EARLY DETECTION OF DIABETIC RETINOPATHY USING LOW RESOLUTION IMAGES

Synopsis of the Ph.D. Thesis

Submitted by

Vidya Sagvekar

(Roll No: 70821016018)

Department of Electronics and Telecommunication Engineering

Supervisor(s):

Dr. Manjusha Joshi (Guide)
&
Dr. Minu Ramakrishnan (Co-Guide)

Submitted in partial fulfilment of the requirements of
DOCTOR OF PHILOSOPHY (Engineering)

MUKESH PATEL SCHOOL OF TECHNOLOGY MANAGEMENT & ENGINEERING
SVKM's NARSEE MONJEE INSTITUTE OF MANAGEMENT STUDIES
(Declared as Deemed-to-be University Under Section 3 of the UGC Act, 1956)
V. L. Mehta Road, Vile Parle (West), Mumbai – 400056

August 2024

Abstract

Diabetes Mellitus (DM) is a widespread chronic condition that can lead to complications such as Diabetic Retinopathy (DR), a major cause of vision impairment in individuals with diabetes. DR often remains undiagnosed until its advanced stages, making early detection essential to prevent irreversible vision loss. This research presents a novel automated framework for the early detection and classification of DR, leveraging advanced image-processing techniques and deep-learning models. This research incorporates both low-resolution images, captured using a smartphone-based device, and high-resolution images, obtained through traditional fundus cameras. Additionally, publicly available dataset images are utilized to develop and evaluate the proposed methods.

The first contribution involves a Chronological Rider Sea Lion Optimization (CRSLO)-based Rider Neural Network (RideNN) for DR classification, with SeGAN employed for accurate lesion segmentation. The second contribution introduces the Hunter-Prey Ladybug Beetle Optimizer Deep Maxout Network (HPLBO_DMN), which integrates K-Net with HPLBO to optimize both lesion segmentation and classification performance. The third contribution features a CAViaR Dung Beetle Optimization-enabled DenseNet (CDBO_DenseNet) model, combining CAViaR and Dung Beetle Optimization (DBO) techniques to enhance feature selection and model robustness, achieving a classification accuracy of 94%.

Advanced techniques such as Deep Joint Segmentation were utilized, achieving a segmentation accuracy of 92.3%. Comparative analysis of segmentation methods demonstrated the superior performance of Deep Joint Segmentation. Inter-rater reliability analysis among ophthalmologists resulted in a Cohen's kappa coefficient of 0.85, indicating strong agreement on lesion identification. Validation results of the third proposed model, CDBO_DenseNet, on both low-resolution and high-resolution images, showed accuracies of 93% and 93.5%, respectively, highlighting its effectiveness and adaptability regardless of image resolution.

Additionally, as the training set percentage increased from 50% to 90%, a consistent decrease in loss was observed across all models. Loss analysis showed that the hybrid model outperformed individual optimizers in minimizing loss, regardless of dataset size. These results highlight the model's strong generalization capability and the effectiveness of the optimization strategies.

Contents

Abstract **i**

1 Introduction **1**
1.1 Diabetic Retinopathy 1
1.2 Differences Between Hypertensive Retinopathy and Diabetic Retinopathy 3
1.3 Techniques of Eye Examination 4
1.3.1 Fundus Images 4
1.3.2 Smartphones-based fundus camera 5
1.4 Lesion segmentation 5
1.5 DR Classification 6
1.6 Techniques used for DR classification 6

2 Literature Survey **8**
2.1 Introduction 8
2.2 Related Work 8
2.2.1 Smartphone-based retina imaging tools 8
2.2.2 Review on Lesion segmentation techniques 9
2.2.3 Validate Smartphone images using Fundus camera images 9
2.2.4 Conventional DR classification techniques with their advantages and restrictions 10
2.3 Research Gaps 11

3 OBJECTIVES **13**

4 CONTRIBUTIONS **14**
4.1 Contribution 1 14
4.2 Contribution 2 (Proposed Method 1) 15
4.3 Contribution 3 (Proposed Method 2) 16
4.4 Contribution 4 (Proposed Method 3) 18

5 MATHEMATICAL BACKGROUND **20**
5.1 Feature Extraction 20

6 SeGAN-Based Lesion Segmentation and Optimized Ridenn for Diabetic Retinopathy Classification 24
6.1 Introduction 24
6.2 Problem Formulation 24
6.3 Proposed Methodology 25
6.4 Experimental Result 32

7 Hybrid hunter-prey Ladybug beetle optimization enabled deep learning for diabetic retinopathy classification 34
7.1 Introduction 34
7.2 Proposed Methodology 35
7.3 Experimental Analysis 47

8 CAViaR Dung Beetle Optimization Enabled Deep Learning for Diabetic Retinopathy Classification 49
8.1 Introduction 49
8.2 Proposed Methodology 49
8.3 Experimental Results 57

9 DATASET 64

10 Conclusion 66

11 Future Scope 68

12 PUBLICATIONS 69

References 70

Chapter 1

Introduction

At the Global level, Diabetic Retinopathy is the essential cause of vision loss in patients affected with diabetes. According to the World Health Organization (WHO), in the year 2030, nearly 79.4 million persons will be affected by DM in India. In contrast, more than 18% of the diabetic population are suffering from DR. Detection in the initial stage is required to minimize the risk of blindness in the case of DR. The traditional fundus camera is employed for DR monitoring; nevertheless, its performance is limited to initial healthcare centers, is expensive, and requires skilled personnel. These days, to tackle the issues of fundus cameras, smartphone-based fundus camera imaging is employed for screening the DR. At present, fundus monitoring is mainly diagnosed by manually evaluating the fundus images, which require ophthalmologists with professional clinical experience. Due to limitations on the availability of ophthalmologists in rural parts of India, smartphone-based fundus imaging provides the facility for optometry to examine the eye condition of diabetic patients regularly to prevent vision loss by early diagnosis, in further assist them in visiting tertiary health care centers. So, DR diagnosis has the foremost practical consequence. Nevertheless, the segmenting lesion of DR is the required work for monitoring, as well as for DR diagnosis.

1.1 Diabetic Retinopathy

In today's world, diabetes is the foremost problem for people. According to the research of WHO, the world prevalence of diabetes is demonstrated as 2.8% in the year 2000 as well as it is expected to increase to 4.4% in the year 2023. The contributing factors to the diabetes increasing rate are the aging population, lack of exercise, and obesity. The rising rate of diabetes worldwide across all age groups is hence increasing the likelihood of problems related to diabetes [1]. The most severe form of diabetes mellitus (DM) in adults of working age is called diabetic retinopathy (DR), and it is the main cause of blindness. Clinical signs of vascular abnormalities in the retina diagnose diabetic retinopathy (DR)

[2]. The increasing amount of diabetes is not only occurring in developing nations, but it also affects non-developing countries. Regrettably, many developed countries are lacking in the essential recording of DR cases. It is demonstrated that more than 75% of people are living with DR in emerging nations. The emergent notion's situations are awful because of insufficient treatment. Nonetheless the health care circumstances in basic countries, people who have diabetes are prone to blindness when compared with people who do not have the illness. DR is considered a silent ailment, for the reason it is only predicted to be realized by patients when the retina changes have developed to the level when treatment becomes difficult, as well as nearly impossible. Depending on when diabetes first appears and how long the condition has been present, different retinopathy prevalences occur [3] injured blood vessels, producing leaks in the blood, the explosion of micro as well as the other fluids, swelling of the retinal tissues, causing blurred sight in both eyes. The DR is progressive if enhancing DR in any person with diabetes is thoroughly connected to the period of poorly managed diabetes. The chance of developing blindness from untreated DR is increased [4].

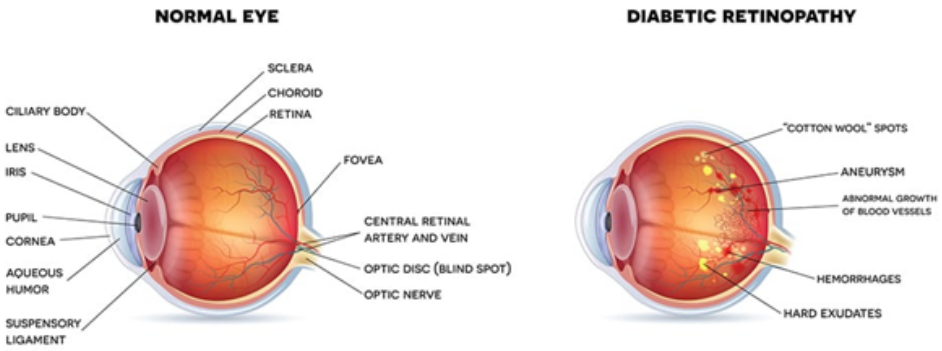

Figure 1.1: (a) Normal eye (b) Different lesions in diabetic retinopathy [[5]

Figure 1.1 (a) shows the normal eye image, while Figure 1.1 (b) represents lesions that cause damage to the retinal blood vessels. This damage can lead to leakage of blood and other fluids, resulting in retinal hemorrhages, microaneurysms, exudates, and cotton wool spots.

Initially, patients with diabetic retinopathy may not experience any noticeable changes in their vision. The earliest signs of the condition involve tiny areas of swelling in the retina's small blood vessels. Once symptoms appear, they may include the following:

- Hazy or double vision
- Observing dark floaters, blank spots, or flashing lights
- Pressure or pain in one or both eyes
- Issues related to peripheral vision

In diabetic retinopathy (DR), the following features provide valuable diagnostic information.

These features should be carefully noted while examining fundus images:

1. **Microaneurysms (MAs):**
    - Small, round dilations in retinal blood vessels.
    - Early signs of diabetic retinopathy (DR).
    - Appear as tiny red dots.
2. **Dot-and-Blot Hemorrhages:**
    - Small, irregularly shaped hemorrhages.
    - Located in the inner retinal layers.
    - Associated with vascular fragility.
3. **Flame-Shaped Hemorrhages:**
    - Linear, flame-like hemorrhages.
    - Found in the nerve fiber layer.
    - Often related to hypertension.
4. **Hard Exudates:**
    - Yellow or white deposits.
    - Accumulate in the outer retinal layers.
    - Result from lipid leakage due to damaged blood vessels.
5. **Cotton Wool Spots:**
    - Soft, fluffy white patches.
    - Indicate retinal nerve fiber layer ischemia.
    - Associated with poor blood flow.

1.2 Differences Between Hypertensive Retinopathy and Diabetic Retinopathy

Hypertensive retinopathy and diabetic retinopathy are two distinct conditions that affect the retina, each with different causes, pathological changes, symptoms, and management strategies. Hypertensive retinopathy is caused by high blood pressure, whereas

diabetic retinopathy results from hyperglycemia associated with diabetes. The pathological changes in hypertensive retinopathy primarily involve larger retinal vessels and can lead to arteriosclerosis, whereas diabetic retinopathy predominantly affects smaller retinal vessels, resulting in microaneurysms and neovascularization. In terms of symptoms and signs, hypertensive retinopathy is characterized by features such as arteriovenous (AV) nicking and optic disc edema. In contrast, diabetic retinopathy presents with microaneurysms, retinal hemorrhages, and neovascularization. Management of these conditions also differs; hypertensive retinopathy is mainly managed by controlling blood pressure, while diabetic retinopathy requires comprehensive management of blood sugar levels, along with controlling blood pressure and lipid levels.

1.3 Techniques of Eye Examination

Early detection of DR is crucial for preventing vision loss, as timely intervention can halt or slow the progression of the disease. Traditionally, the diagnosis of DR is performed through comprehensive eye examinations involving techniques such as pupil dilation, fundus photography, optical coherence tomography (OCT), and fluorescein angiography. These methods, while effective, are resource-intensive, time-consuming, and require specialized equipment and expertise, making them less accessible, especially in low-resource settings.

1.3.1 Fundus Images

The usage of fundus images for timely screening of eye illnesses is of the highest medical importance. The projection from the fundus image is the fundus capture from the monocular 2D plane camera. The fundus images are obtained cost-effectively and non-invasive, constraining them for large-scale screening. In the fundus image, many different Biomarkers are attained, they are blood vessels, fovea, macula, optic cup, and optic disc as well as some of the DR-related lesions, such as soft exudates, hard exudates, hemorrhages, and microaneurysms. Fundus images are useful in the diagnosis of retinopathy of prematurity, diabetic macular edema, cataracts, and glaucoma, among other eye conditions. [[6]]. In the eyes, fundus images are employed to folder or record the states as well as to evaluate the variations of healthy conditions. Imaging methods may focus on the structure of the retina, a specific functional feature (such as fluorescein angiography or color fundus photography), or a correlation between the two [12]. Fundus cameras (non-mydriatic as well as mydriatic): The fundus images are captured with the help of a fundus camera, which can be mydriatic or non-mydriatic. Mydriatic, which uses a bright light flash and pupil dilation. More sophisticated non-mydriatic requires no pupil dilation and only uses a low-intensity flashlight [7].

1.3.2 Smartphones-based fundus camera

The smartphone is a user-friendly, portable, and handheld device that can be effectively used for capturing fundus images. The smartphone can be employed sufficiently by young children, and adults operation also for clinical purposes. For vision tests in groups, smartphones are more practical for corresponding camps. These can also be utilized in inadequate circumstances, such as insufficient lighting in an accident scene or an emergency department when it is difficult to employ an indirect ophthalmoscope or fundus camera due to their high cost [4]. The smartphone price, together with 20D lenses, is around 500$. But a fundus normal camera costs around 10 thousand US dollars, which is more expensive compared with the smartphone cost. At present, the indirect ophthalmoscope and fundus camera are obtainable for usage; moreover, employing that equipment is no longer possible because of insufficient electricity in rural areas. The application in smartphones is portable, lightweight, equipment is easily available, and cost-effective. Advancement in mobile technology, such as video/photo capturing, storage of data, and sharing images and internet connectivity, which provides an easy way of consultation, proves it the finest device for retinal ophthalmoscopes [7]. The aim of the smartphone in the medical image field has expanded to encompass a huge assortment of purposes, like predicting retinal diseases and monitoring. The app on smartphones is employed by physicians for medical practices to diagnose and monitor patient images. Similarly, smartphones provide a platform for ophthalmologists to view, record, and share results in the form of images in a more economical way [4]. Smartphones used for capturing fundus images can be branded into numerous types based on the specification, such as smartphones with a 20D lens model, D-Eye, oDocs Fundus, iExaminer, Cell Scope Retina, PEEK, and JaizRetiCam.

1.4 Lesion segmentation

Lesion segmentation is an essential step in the screening and diagnosis of DR. It also establishes the framework for the subsequent severity grading of DR [8]. The problem may occur in the process of segmenting the lesion. Some of the issues are as follows: a) the lesion proportion in a fundus image with high resolution is feasible for noise confusion because it is tiny. b) The next issue is that lesion structures are complex, as well as the brightness, color, size, and shape, which are different for each lesion. c) the presence of false positive outcomes due to the reason of texture, contour, and color of retinal tissues. d) The retinal image appearance is different because of the light source intensity and the camera system. Lesion segmentation becomes more challenging when dealing with fundus images of patients from different racial backgrounds due to the wide variations in tissue pigmentation levels [7]. Precisely, comparing the retina images of yellow people

and white people, melanocytes are more present in black people. Compared to the light-colored retina, the darker pigmentation of the retina can hazy, subtle abnormalities, making them harder to detect [9]

1.5 DR Classification

DR has two primary stages: proliferative DR (PDR) and non-proliferative DR (NPDR) [2]. The DR preliminaries stage is represented as NPDR, wherein capillary occlusion and enhanced vascular permeability occur. Initial stages, patients may be asymptomatic. At the present stage, fundus imaging can be used to identify retinal diseases, including micro-aneurysms, hemorrhages, and hard exudates [10]. DR advanced stage is PDR, which considers using neovascularization. By this stage, vision impairment is experienced by the patient because, in the vitreous vessels, bleeding occurs. Proliferative Diabetic Retinopathy is the main reason for loss of vision in the patient. DME can occur at any stage during DR and result in blurred vision and reduced visual acuity [2]. The quantity of retinal lesions varies among DR grading systems, as different NPDR and PDR grading criteria are listed in Table 1.1 based on the International Clinical DR Severity Scale.

Table 1.1: Diabetic Retinopathy Severity Levels and Associated Findings

Disease Severity Level	**Description**	**Findings Observable on Dilated Ophthalmoscopy**
Grade 0	No retinopathy	No abnormalities
Grade 1	Mild NPDR	Microaneurysms only
Grade 2	Moderate NPDR	Between just microaneurysms and severe NPDR
Grade 3	Severe NPDR	More than 20 intraretinal hemorrhages in each of 4 quadrants; venous beading in more than 2 quadrants
Grade 4	PDR	Neovascularization

1.6 Techniques used for DR classification

At present, the enhancement of Deep learning (DL) is majorly used in the medical treatment field. Better results were achieved by Deep Neural Networks (DNNs) in many domains, such as voice and computer vision.

The DNN has attained performance that surpasses human beings, for example, in large-scale visual recognition and face recognition. In diagnosing DR, the DNN has made

many improvements and has attracted much interest. Currently, a lot of effort needs to be made because clinical use of automatic DR detection systems is currently unavailable despite the numerous advancements that have been accomplished [11]. A Cascade Attentive Refine Net (CAR_Net) for segmenting numerous lesions. For automated and precise multi-lesion DR segmentation, CAR_Net uses an encoder-decoder architecture and end-to-end training[9]. The Deep Learning model developed based on the Artificial Neural Network (ANN) is an improved stream of Machine Learning (ML) to study abstraction using data in a multi-layered fashion. It demonstrated that Deep Learning models are extremely effective in the assorted range of biomedical imaging analytics and computer[12]. The Deep Convolutional Neural Network (CNN) is one of the successful strategies in medical image classification and is used by researchers to classify medical [13].

Chapter 2

Literature Survey

2.1 Introduction

DR is an eye condition that may result in blindness and visual impairment. Timely diagnosis and treatment are crucial for preventing potential complications. The classification of DR is crucial for documenting a patient's disease status. In this chapter, we discuss several methods used for DR that are broadly categorized as Deep Learning, Machine Learning, Transfer Learning (TL), and other Artificial Intelligence (AI) based techniques. The categorized techniques and the survey papers that have adopted these techniques are discussed below. Finally, the chapter concludes with a summary of research gaps and difficulties encountered by the methods.

2.2 Related Work

2.2.1 Smartphone-based retina imaging tools

The author of this study compares the image quality and DR classification capabilities of several smartphone-based portable retinal imaging devices that are currently on the market by utilizing a deep learning framework [6]. For iNview, iExaminer, D-Eye, Peek Retina, and iNview photos, the overall classification accuracy of smartphone-based systems is ranked as follows: 75%, 61%, 62%, 69%. As the field of view of retinal systems varies based on the type of retinal devices. Field view of iNview being the largest and iExaminer the smallest—the author noticed that the network DR detection performance diminishes. Retinal imaging with smartphone fundus photography is an affordable, practical, portable, and efficient technique. Smartphone-based camera apps are made specifically to be used easily by beginning users, and users can become proficient in this method with practice. This retinal imaging method can be used for artificial intelligence (AI), patient diagnosis, and teaching by educating young ophthalmology residents and ophthalmic primary caretakers [14]. The image quality obtained from this

new smartphone-based fundus imaging device was sufficient to document a variety of fundus diseases, including only mild discoveries, although it was still not as good as that of conventional fundus imaging. A portable substitute for fundus documentation in clinical practice could be high-quality smartphone-based fundus imaging [80][15]. In 2017, Martínez-Aguirre stated that the primary instruments used in ophthalmology to examine and capture images of the eye fundus are fundus cameras. It is also called a direct and indirect ophthalmoscope. Conventional fundus cameras provide high-quality images, but they are expensive, large, reliant on a technician, and space-consuming. Smartphones with built-in high-definition cameras are used as fundus imaging devices, giving users the opportunity to take fundus images in difficult-to-reach locations like rural areas without access to healthcare, developing-nation humanitarian projects, surgical centers, wards in hospitals, and medical facilities [16]. Fundus imaging is carried out simply by Haddock et al. using a smartphone, a 20D lens, and either a Koeppe lens or not. The device functions as a kind of indirect ophthalmoscope. HD videos of the retina were captured and recorded using the built-in app, and high-quality images were extracted from the video clips [17].

2.2.2 Review on Lesion segmentation techniques

Zhou et. al. propose a new model named FFU-Net which is suitable for lesion segmentation of diabetic retinopathy [8]. Here is a convolutional layer to reduce the spatial loss of the fundus image in place of the pooling layer. To extract multiscale features and expedite feature fusion with a splitting operation, the MSFF block is implemented. Through the usage of the CCA module, FFU-Net integrates data from lower-resolution decoders with shared attention weights and skip connections. The public benchmark IDRiD is used for user studies. FFU-Net improves the segmentation performance based on metrics SEN, IOU, and DICE, respectively. Guo et.al.[15][9] proposed a cascading attentive RefineNet (CARNet) for precise and automated multi-lesion segmentation of diabetic retinopathy. With the IDRiD, E-ophtha, and DDR data sets, the author assessed the segmentation performance of the proposed CARNet. Qiao, L. Zhu[18][18] et al. proposed a semantic segmentation method to categorize fundus images as either DR or No DR. Semantic segmentation segments the image's pixels according to their shared semantics to detect the microaneurysm characteristic. Ophthalmologists will be able to categorize fundus images as moderate or severe NPDR by using this automated technique.

2.2.3 Validate Smartphone images using Fundus camera images

Ramachandra et. al. perform validation of smartphone-based retinal images for Diabetic retinopathy screening. Used smartphone-based retinal imaging system "FOP". Sensitivity and Specificity performance metrics were evaluated to identify the effectiveness of

the screening tool. The sensitivity and specificity of retinal imaging using FOP for any DR were very high (92.7% and 98.4%) and for STDR was also high (87.9% and 94.9%), respectively [19].

2.2.4 Conventional DR classification techniques with their advantages and restrictions

In 2018, Wan S., et al.[20] introduced Deep CNN for DR detection. The input image experienced pre-processing to enhance image quality, which enhances classification performance. Despite having a lower computing cost, this method's model classification performance is inaccurate because of the huge number of tracing parameters and a small amount of training data. In 2019, Fatema T. Johora., et al.[21] developed a DR detection model utilizing weighted decision tree-based classification and PCA-SIFT-based feature extraction. This technique was more compact and unique for both controlled and real-world circumstances, resulting in significant improvements in matching accuracy.

However, the proper work is done with the help of high-resolution figures. In 2020, Gayathri S., et al.[22] developed a lightweight CNN approach for DR classification from fundus images. Using CNN, feature extraction was done using a fundus image as input. The classification of DR was then done using ML classifiers. This approach meaningfully decreased the processing time required for the classification procedure, but it was unable to analyze all retinal illnesses. Saranya P. and Prabakaran S., [10] developed a model of the CNN technique for DR identification in retinal fundus images to identify other types of retinal disorders. To perform extra processing after receiving the retinal image, the optic disc was segmented. Pre-processing and CNN-based categorization were applied to the input image. This technique aids clinical practitioners in early disease prognosis. Nevertheless, this method did not allow for the identification of the stages of DR during diagnosis. Asia, A.O., et al. [3] developed a CNN model to distinguish the stages of DR. This method carefully assessed various CNNs while taking into account the criteria that were used to identify the severity of the sickness, and it supported the clinical applicability of those CNNs. However, the database is small, noisy, and very vast in size. Deep networks are thus a little challenging to train. Araújo, T., developed the NeoVessel (NV) generation algorithm introduced in [11], for improving proliferative DR detection. Multiple datasets were used to test this approach, enhancing the model's performance. However, part of the PDR images that do not contain NVs are still not recognized, as the model has not been trained to recognize pre-retinal fibrosis or pre-retinal hemorrhages. Kathiresan S., et al. [8][23] presented a Synergic Deep Learning (SDL) approach for DR detection and classification to improve the detection outcome. The input image was pre-processed to eliminate noise and other artifacts, and the segmentation procedure was carried out using histogram-based segmentation.

Finally, the SDL technique was used to categorize DR classification into different severity levels. This methodology shortens the time it takes to complete a classification, even if this model didn't include filtering procedures for improving image quality before processing. Laxmi Math and Ruksar Fatima [24] used the segment-based learning approach for detecting DR. It can learn classifiers mutually and features from the data. However, it may not be fully compatible or transferable to task. In 2021, Qureshi, I., et al.[6] devised an Active deep-learning Convolutional Neural network (ADL-CNN) for DR detection. This technique provided useful prognostication masks to help healthcare professionals with the essential duty of labeling eye samples and segmenting areas of interest within fundus images to grade severity levels of DR. However, certain photos on the Proliferative DR, severe, and moderate DR severity categories could not be classified using this technique. In 2022, Kshirsagar, P.R., et al.[7] established a deep-learning neural network to identify DR in the retina's blood vessel. Using a neural network-based architecture, this technique automatically segments retinal hemorrhages in a fundus image. Here, entropy, energy, contrast, or homogeneity was not taken into account.

2.3 Research Gaps

This work addresses the following research gaps associated with the detection of DR

1. Uneven illumination in fundus images is a common challenge that can hinder the performance of deep-learning models for detecting diabetic retinopathy and other retinal diseases. Addressing this issue involves a combination of preprocessing techniques to normalize the lighting and improve the image quality.

2. Traditional image processing techniques have limitations in detecting subtle lesions in diabetic retinopathy due to their sensitivity to image quality variations, lack of generalization, reliance on manual feature extraction, and poor integration with clinical workflows. Addressing these gaps requires the development of more robust, adaptive, and comprehensive methods, potentially leveraging advances in deep learning and multi-modality data integration to improve the early detection and diagnosis of diabetic retinopathy.

3. Existing methods often fail to classify diabetic retinopathy (DR) stages accurately, particularly in the early stages where symptoms are subtle. There is a shortage of large, diverse, and well-annotated datasets. Additionally, existing datasets often suffer from class imbalance, with fewer examples of certain stages of DR. Current models struggle to accurately identify and differentiate between the various stages of DR, especially the early and subtle signs of disease progression. Models trained on specific datasets may not generalize well to different populations or imaging devices, limiting their applicability in diverse clinical settings. Addressing these research gaps involves developing more accurate and reliable classification models, creating larger and more diverse datasets,

and improving the identification of disease stages.

4. To address the research gaps in handling smartphone images with poor resolution for diabetic retinopathy detection using deep learning, efforts should focus on improving image quality, enhancing model robustness and generalization, creating diverse and comprehensive datasets, developing advanced feature extraction methods, ensuring real-time processing capabilities, and conducting clinical validation. By addressing these gaps, the deployment of effective and reliable DR detection models on smartphones can be realized, particularly benefiting low-resource settings where access to professional fundus cameras is limited.

Chapter 3

OBJECTIVES

To fulfill the research gaps, the following objectives were set for this work:

1. To investigate techniques for lesion segmentation, deploy them to identify early lesions, and evaluate their performance.

2. To study and evaluate the performance of deep learning techniques for the classification of diabetic retinopathy into No DR, mild NPDR, moderate NPDR, severe NPDR, and PDR.

3. To develop hybrid optimization techniques to improve the performance of deep learning models used for the classification of diabetic retinopathy.

4. To develop models by integrating studied methods for lesion segmentation and DR classification with the proposed hybrid optimizer, including extracting additional features from the retina image and comparing their performance.

5. Performance comparison between low-resolution and high-resolution images using metrics like accuracy, sensitivity, and specificity.

Chapter 4

CONTRIBUTIONS

After identifying the research gaps and limitations of existing techniques, this study proposes the following methods, with the contributions detailed below:

4.1 Contribution 1

1. Survey and Analysis: Conducted a comprehensive survey of various techniques for fundus image pre-processing, lesion segmentation, and classification methods used in the detection of diabetic retinopathy. Additionally, reviewed smartphone-based retina imaging devices and their validation techniques.

2. Data Collection and Dataset Creation: Regularly visited the Ophthalmology department at Somaiya Hospital to capture and collect smartphone-based retina images and fundus camera images of diabetic patients. As a result, a dataset was created comprising 120 real-time fundus images captured using a traditional fundus camera and approximately 84 smartphone-based images. Also performed a survey and downloaded publicly available datasets of fundus images.

3. Image Quality and Validation: Conducted a validation of smartphone-based camera images against fundus camera images in terms of image quality. Evaluated the sensitivity, specificity, and Cohen's Kappa Agreement to assess the level of agreement between the two imaging modalities.

4. Ground Truth Collection and Inter-Rater Reliability: Collected ground truth data from two ophthalmologists at Somaiya Hospital and performed inter-rater reliability analysis. The inter-rater reliability analysis yielded a Cohen's Kappa coefficient of 0.85, indicating strong agreement among the ophthalmologists regarding the identification of diabetic retinopathy lesions.

5. Clinical Validation: Executed clinical validation between smartphone-based imaging devices and conventional fundus cameras using Cohen's Kappa agreement, achieving a coefficient of 0.78, which demonstrates substantial agreement between the two modalities.

Objectives Achieved:

1. Dataset development: Successfully built a dataset of fundus images and smartphone-based retina images for the detection of diabetic retinopathy.

2. Validation of smartphone imaging: Achieved a Cohen's Kappa coefficient of 0.78, showing substantial agreement between smartphone-based retina imaging devices and conventional fundus cameras.

3. Inter-rater reliability: Conducted inter-rater reliability analysis with a Cohen's Kappa coefficient of 0.85, indicating strong agreement among ophthalmologists' assessments of diabetic retinopathy lesions.

4.2 Contribution 2 (Proposed Method 1)

1. Development of a Novel Hybrid Deep Learning Model (CRSLO-based RideNN): Introduced a novel hybrid deep learning model, CRSLO-based RideNN (Chronological Rider Sea Lion Optimization-based Rider Neural Network), specifically designed to improve the accuracy and efficiency of diabetic retinopathy (DR) detection and classification. This model utilizes a hybrid optimization approach to enhance both the detection process and the overall performance.

2. Preprocessing and ROI Extraction: Applied advanced preprocessing techniques, including Median Filtering and Region-of-Interest (ROI) extraction, to enhance image quality and focus on critical areas for detecting DR-related features.

3. Multiclass Lesion Segmentation: Used a deep learning-based segmentation technique, the Speech Enhancement Generative Adversarial Network (SeGAN), to achieve simultaneous segmentation of multiple DR-related lesions, such as microaneurysms, hemorrhages, hard exudates, and soft exudates. This approach enables comprehensive lesion analysis within retinal images.

4. Blood Vessel Segmentation and Classification: Developed a method for precise blood vessel segmentation using sparking processing techniques. Further classified arteries and veins using a Support Vector Machine (SVM), providing detailed vascular analysis crucial for accurate DR detection.

5. Feature Extraction and DR Classification: Extracted essential features from segmented images for effective DR classification. Employed the RideNN model, optimized using the novel CRSLO hybrid approach, which combines the strengths of the Chronological concept, Rider Optimization Algorithm (ROA), and Sea Lion Optimization (SLnO). This hybrid optimization enhances model performance by improving convergence speed, stability, and accuracy.

6. Comprehensive Evaluation on Benchmark Datasets: Evaluated the CRSLO-based RideNN model using standard benchmark datasets, such as DIARETDB1 and DDR. Results demonstrated superior accuracy and computational efficiency compared to existing methods, highlighting the effectiveness of the hybrid optimization approach.

Objectives Achieved:

1. Achieved Advanced Multiclass Lesion Segmentation: The model effectively delineates different lesion types (such as microaneurysms, hemorrhages, hard exudates, and soft exudates) from fundus images, ensuring precise localization of pathological regions.

2. Extract the features from segmented images, such as Local Vector Pattern (LVP), Local Gabor Binary Patterns (LGBP), Perimeter, Kurtosis, Entropy, Eccentricity, Area, Color histogram, Artery to Veins ratio for classifying DR effectively.

3. Developed a Robust and Accurate DR Classification Model: Created a state-of-the-art DR classification model using the novel CRSLO-based RideNN approach, demonstrating significant improvements in accuracy, computational efficiency, and adaptability to diverse clinical datasets.

4. Novel Hybrid Optimization Approach: The integration of three optimization techniques (Rider Optimization, Chronological Concept, and Sea Lion Optimization) is a novel approach that significantly improves convergence speed, reduces the chances of getting trapped in local minima, and enhances overall classification accuracy.

4.3 Contribution 3 (Proposed Method 2)

1. Introduction of a Novel Hybrid Optimized Deep Learning Model (HPLBO_DMN): This work proposes a novel hybrid optimized deep learning model, the Hunter-Prey Ladybug Beetle Optimizer Deep Maxout Network (HPLBO_DMN), designed to enhance the accuracy and robustness of diabetic retinopathy (DR) classification.

The model leverages advanced hybrid optimization techniques to improve both the segmentation and classification processes.

2. Advanced Image Preprocessing and Enhancement: Implemented a comprehensive image preprocessing pipeline, including the use of the Kalman filter and Region of Interest (ROI) extraction. The Kalman filter adaptively estimates and corrects spatially varying illumination in fundus images, thereby enhancing the overall image quality. Subsequently, a wavelet transform using the Meyer wavelet is applied to improve the edge sharpness index, which is critical for accurate lesion detection.

3. Novel Lesion Segmentation Method (K-Net_HPLBO): Developed a novel lesion segmentation approach using K-Net, a deep learning-based network that is trained with the proposed Hunter-Prey Ladybug Beetle Optimizer (HPLBO). This segmentation method enables precise detection of lesions associated with DR, such as microaneurysms, hemorrhages, and exudates, enhancing the model's ability to identify and differentiate between DR severity stages.

4. Automatic Artery/Vein Classification (AVNet): Use an Automatic Artery/Vein Classification Network (AVNet) that simultaneously classifies arteries and veins from the wavelet-transformed images. This step is crucial for detailed vascular analysis, which plays an important role in DR diagnosis.

5. Feature Extraction and DR Classification: The feature extraction phase leverages the output of the AVNet to obtain meaningful features from the retinal images. These features are then fed into a Deep Maxout Network (DMN) for DR classification. The DMN is trained using the proposed HPLBO, which is a hybrid optimization method combining the Hunter–Prey Optimizer (HPO) and the Ladybug Beetle Optimization (LBO). This hybrid optimization technique ensures optimal training of the DMN, leading to effective classification into DR severity grades: normal, mild NPDR, moderate NPDR, severe NPDR, and proliferative DR.

6. Extensive Evaluation on Multiple Datasets: The proposed HPLBO_DMN model is extensively evaluated on different datasets, including STARE, Messidor, DRIVE, DIARETDB1, and a real-time dataset comprising both high-resolution fundus camera images and low-resolution smartphone images. Novelty of the Contribution: The novelty of this research lies in the innovative use of a hybrid optimization technique, the Hunter-Prey Ladybug Beetle Optimizer (HPLBO), which combines the strengths of two distinct optimization algorithms: the Hunter–Prey Optimizer (HPO) and the Ladybug Beetle Optimization (LBO). This hybrid approach enhances the model's training process by improving convergence speed, stability, and

accuracy. The HPLBO effectively balances exploration and exploitation during training, allowing the model to avoid local minima and achieve superior performance in DR classification. This novel optimization strategy is particularly effective in handling the complex nature of retinal images and contributes significantly to the overall robustness and accuracy of the proposed model.

Objectives Achieved with This Contribution:

1. Accurate Lesion Segmentation Using K-Net: Implemented K-Net for lesion segmentation, trained with the proposed Hunter-Prey Ladybug Beetle Optimizer (HPLBO), enhancing segmentation accuracy and enabling precise identification of DR lesions.

2. Extract the features from segmented images, such as Local Vector Pattern (LVP), Local Gabor Binary Patterns (LGBP), Perimeter, Kurtosis, Entropy, Eccentricity, Area, Color histogram, Artery to Veins ratio for classifying DR effectively.

3. Development of a Robust and Accurate DR Classification Model: Successfully developed the HPLBO_DMN model, which demonstrates superior accuracy and robustness in classifying DR severity levels compared to existing models.

4. Improved Model Training with Hybrid Optimization: Demonstrated the effectiveness of the novel hybrid optimization technique (HPLBO) in training deep learning models, achieving better convergence, reduced training time, and higher accuracy in DR classification tasks.

5. Validation across Multiple Datasets: Validated the performance of the proposed model on multiple datasets, including both high-resolution and low-resolution images, showcasing its adaptability and scalability for real-world clinical applications.

4.4 Contribution 4 (Proposed Method 3)

1. Introduction of a Novel method named CAViaR Dung Beetle Optimization-DenseNet (CDBO_DenseNet) : The novelty of this approach lies in the integration of hybrid optimization techniques—CAViaR and Dung Beetle Optimization—with the DenseNet architecture. This combination significantly enhances the efficiency of feature selection and classification tasks within the model.

2. Pre-processing: The process begins with a pre-processing module, where a Mean filter is employed to improve image quality. This enhancement is crucial for ensuring that the subsequent steps receive a cleaner and more defined input. Following pre-processing, lesion segmentation is carried out using Deep Joint Segmentation, a

technique that accurately identifies and isolates pathological regions within retinal images.

3. Next, the classification of arteries and veins is performed using AVNet, which plays a pivotal role in refining the vascular analysis within the images. The feature extraction process is then optimized by the hybrid CDBO technique, allowing the model to focus on extracting the most relevant and essential features, thereby improving its overall classification accuracy.

4. The CDBO_DenseNet model was evaluated on seven different datasets: STARE, Messidor, DRIVE, DIARETDB1, IDRID, DIARETDB0, and a real-time dataset that includes both high-resolution images and low-resolution smartphone images. The validation of the proposed model was performed using a training set (varying from 50% to 90%) and K-fold (varying K value 6 to 9) cross-validation technique. Also, Loss analysis was performed on individual and proposed hybrid optimizers, demonstrating their capability to effectively handle different image sources. Additionally, a confusion matrix was plotted to visually represent the classification outcomes across five distinct classes: Proliferative, Severe NPDR, Moderate NPDR, Mild NPDR, and Normal. The proposed model was compared with existing methods and earlier proposed models to evaluate its performance.

Objectives Achieved with This Contribution:

1. Enhanced lesion segmentation accuracy by employing the Deep Joint Segmentation technique, outperforming previously proposed methods.

2. The hybrid optimization approach led to enhanced feature selection, significantly improving DR classification accuracy.

3. The model showed robust performance across diverse datasets, including low-resolution images, confirming its adaptability.

4. Through cross-validation, the model demonstrated improved generalization capabilities, making it suitable for real-world applications.

Chapter 5

MATHEMATICAL BACKGROUND

Introduction This work combines various engineering techniques for feature extraction, optimization, and classification, so a thorough explanation of these methods is provided in this chapter.

5.1 Feature Extraction

In this case, the artery-to-vein ratio, color histogram, eccentricity, area, kurtosis, perimeter, local variance projection (LVP), entropy, and local gabor binary pattern (LGBP) are among the significant features that were extracted.

Artery-to-vein ratio:

Each venule and arteriole value is combined to calculate the artery-to-vein ratio [25].

$$\textit{Arterioles:} \quad \Omega_t = \sqrt{0.87\Omega_j^2 + 1.01\Omega_e^2 - 0.22\Omega_e - 10.76} \tag{5.1}$$

$$\textit{Venules:} \quad \Omega_v = \sqrt{0.72\Omega_j^2 + 0.91\Omega_e^2 + 450.05} \tag{5.2}$$

Here, the caliber of the small branch and large branch is signified as C_j and C_e, respectively, and the caliber of the trunk vessel is denoted as C_t and C_v. The artery-to-vein ratio is indicated as AVR l_1.

Area:

The area feature is referred to as and the total number of pixels in the fundus image that are covered by a particular lesion (such as a hemorrhage, microaneurysm, or exudate). This test aids in determining the degree and course of diabetic retinopathy.

Entropy:

The amount of information contained in high-information pixels, especially in areas close to edges and corners, is mainly measured using entropy. The feature of entropy is signified as l_3.

Eccentricity:

Eccentricity is measured as the major axis measurement divided by the separation between the ellipse centers. Line segments are given the value 1, while circles are given the value 0, which is specified as l_4.

Kurtosis:

Kurtosis, also known as the sharpness of the curve in the frequency distribution, offers information on peakness relative to normal distribution. Furthermore, the feature of kurtosis is articulated as l_5.

Perimeter:

This feature is defined as the sum of the boundary pixels that constitute a single segment. It is considered based on the distance between each boundary region and its surrounding pixel pair. This estimated value of the feature depends on the length of the image and breadth, which are determined by ,

$$perimeter = 2 \times l \times v \tag{5.3}$$

Here, the length of the image is represented as , the width of the image is indicated as and the perimeter feature is denoted as l_6 .

Color histogram:

The colour histogram [26] defines the splitting of the colour composition present in the categorized images of the arteries and veins. To represent the colour histogram of the entire image when the colour space is large, the pixels are separated into bins. Using the histogram, it is demonstrated how the image affects the intensity probability dispersion. The histogram probability is provided by,

$$X(f) = \frac{H_f}{m} \tag{5.4}$$

Here, the entire pixels at the grey level are denoted as H_f, and the overall count of pixels is represented as m. The feature of the color histogram is signified as l_7.

LGBP:

The primary objective of LGBP [27] is to modify the LBP (Local Binary Pattern) operator on Gabor input pictures. In this case, LBP histogram characteristics are removed for each Gabor image. The histogram features on each image are altered to create the LGBP feature. Due to the breakdown of Gabor, LGBP features are larger than LBP features. The LGBP operator also reflects the combination of LBP feature extraction and Gabor filtering. The binary format used to specify the threshold value in a picture is provided by,

$$\mathcal{G}(S_k - S_j) = \begin{cases} 1, & S_k \geq S_j \\ 0, & S_k < S_j \end{cases} \tag{5.5}$$

Here, are LBP operator and is constant and . Here, the LGBP feature is expressed as l_8 .

LVP:

This feature [28] calculates the spacing between the targets and neighboring pixels in various directions and from various distances to build the vector representation of all pixels. The Comparative Space Transform (CST) is applied to express pairwise vector orientation to extract micro-patterns. Features are streamlined and extra code is removed by the LVP feature. The following is the expression of the LVP feature:

$$\Re_{e,v}(R_x) = (N(R_{e,v}) - N(R_x)) \tag{5.6}$$

$$\begin{aligned} LVP_{M,K,e}(R_x) = \Big\{ & L_5\big(J_{e,v}(R_x, K), J_{e+45^\circ,v}(R_x), J_{e,v}(R_x), J_{e+45^\circ,v}(R_x)\big), \\ & L_5\big(J_{e,v}(R_x, K), J_{e+45^\circ,v}(R_x), J_{e,v}(R_x), J_{e+45^\circ,v}(R_x)\big), \ldots, \\ & L_5\big(J_{e,v}(R_x, K), J_{e+45^\circ,v}(R_x), J_{e,v}(R_x), J_{e+45^\circ,v}(R_x)\big) \Big\}_{n=1,2,\ldots,M,K-1} \end{aligned} \tag{5.7}$$

$$\begin{aligned} & L_5\big[J_{f_e,v}(R_x, K),\ J_{e+45^\circ,v}(R_x, K),\ J_{f_e,v}(R_x),\ J_{e+45^\circ,v}(R_x)\big] \\ & = \begin{cases} 1, & \text{if } J_{e+45^\circ,v}(R_x, K) - \left(\dfrac{J_{e+45^\circ,v}(R_x)}{J_{f_e,v}(R_x)} \times J_{f_e,v}(R_x)\right) \geq 0, \\ 0, & \text{else.} \end{cases} \end{aligned} \tag{5.8}$$

In this instance, $\mathcal{V}$ stands for separation between the target pixel and its neighbors, whereas L_5 shows a transformation ratio. Furthermore, $J_{e,v}(R_t)$ represents the direction vector at the target pixel. Moreover, the LVP feature is denoted as ℓ_g.

As a result, all of the acquired features are united to form a feature vector, which is defined as,

$$\Psi_\tau = \{\ell_1, \ell_2, \ldots, \ell_g\} \tag{5.9}$$

A further DR classification method uses the resulting feature vector.

5.3 Performance Evaluation Metrics

5.3.1 Accuracy

Accuracy: Accuracy is the ratio of the output that can be accurately recognized while taking into account the entire amount of the brain image.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{5.10}$$

5.3.2 Sensitivity

Sensitivity: In the DR classification, sensitivity is determined to predict the accuracy rate of true negatives, and it is denoted by,

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100 \tag{5.11}$$

5.3.3 Specificity

Specificity: Specificity is calculated to determine the actual positive classification accuracy rate and is signified by,

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100 \tag{5.12}$$

Where, TP indicates true positive, FP denotes false positive, TN states true negative, and FN states false negative.

Chapter 6

SeGAN-Based Lesion Segmentation and Optimized Ridenn for Diabetic Retinopathy Classification

6.1 Introduction

The premature identification and precise treatment help to avoid vision loss for DR patients. Therefore, it is more requisite to categorize the severity grades of DR for recommendation of effective treatment. Here Proposed Chronological Rider Sea Lion Optimization (CRSLO)-based deep learning technique is devised for DR classification.

6.2 Problem Formulation

In existing methods, traditional thresholding is used for lesion segmentation. These methods are solely based on pixel intensity values, it doesn't consider other relationships between the pixels and also does not perform well in hazy view, uneven lighting, or illumination, Single threshold is also a limitation that may lead to inaccurate segmentation. Also, it does not support multiclass segmentation. Hence deep learning techniques can be used instead of traditional methods of segmentation. In the proposed methodology the lesion segmentation is performed to segment lesion regions as Microaneurysms, Hemorrhages, Hard Exudates, and Soft Exudates by Speech enhancement generative adversarial network. Then, the blood vessel segmentation is achieved by sparking processing, and artery and vein classification is done based on a Support Vector Machine (SVM) for classifying artery and vein regions. Once the artery and vein are classified, feature extraction is performed for extracting features, such as Artery to vein ratio, color histogram, area, eccentricity, entropy, kurtosis, perimeter, and Local Gabor Binary Pattern (LGBP). Then, the effective DR detection and classification are carried out by Rider Optimization

Algorithm-Based Neural Network (RideNN). Moreover, the designed CRSLO is used to train the RideNN, which is the hybridization Chronological concept, Rider Optimization Algorithm (ROA), and Sea Lion Optimization (SLnO) for obtaining better performance. The introduced approach achieved better accuracy, sensitivity, and specificity of 92.18%, 93.04%, and 91.54%. In addition, the employed SeGAN model achieved improved segmentation accuracy of 90.30%.

6.3 Proposed Methodology

Input Image:

Images are taken from the prescribed dataset.

Pre-Processing :

Images are pre-processed using the median filter in order to eliminate the noises and outliers available in the images thus, image quality is highly increased.

a) The median filter for the input image is given by,

$$B(M,N) = \underset{S_{E_{MN}}}{\text{median}} \{o(S_r)\} \tag{6.1}$$

where, E_{MN} specifies the neighborhood sub-image window of dimension $M \times N$, and the result of the pre-processing process is denoted as Q_r.

b) ROI Extraction: ROI extraction [29] process is indispensable to mining the specific image computing difficulties are effectively decreased.
ROI is detected by considering the pixel intensity values from the pre-processed image. The output of ROI extraction is specified as Y_r.

Wavelet Transform:

Discrete Wavelet transform, HL wavelet coefficient has more precise information regarding the image for the lesion segmentation.

Lesion Segmentation Using SeGAN:

Here, the output of the wavelet transform G_r is given to SeGAN [30], in which the lesion region is segmented. Generally, the SeGAN model effectively decreases the quantity of training parameters, and thus processing time is decreased; hence, this structure is utilized for the lesion segmentation process.

Here, the SeGAN segments the lesion as four types, including hemorrhages, microaneurysms, hard exudates, and soft exudates. The output of the SeGAN is specified as F_r.

Artery and Vein Segmentation Using Sparking Process

Subsequently, the wavelet transformed output Gr is considered as input for the artery and vein segmentation process. Here, the sparking process [31] is applied to the image Gr to segment the artery and vein. The maximal threshold is selected as the optimum threshold, which generates artery and vein-segmented images. The output of the artery and vein segmentation process is denoted as Vr.

Artery and Vein Classification Based on SVM

SVM effectively classifies the vein and artery regions, which is indicated as Hr.

Feature Extraction

Then, the essential features are extracted from the artery and vein classification output Hr for the further classification process. Here, the significant features, including Artery to vein ratio, color histogram, area, eccentricity, entropy, kurtosis, perimeter, and LGBP are extracted.

Diabetic Retinopathy Classification Using Developed Chronological Rider Sea Lion Optimization-Based RideNN

After generating the feature vector, the classification of DR is achieved by CRSLO trained RideNN classifier from the formulated feature vector Br.

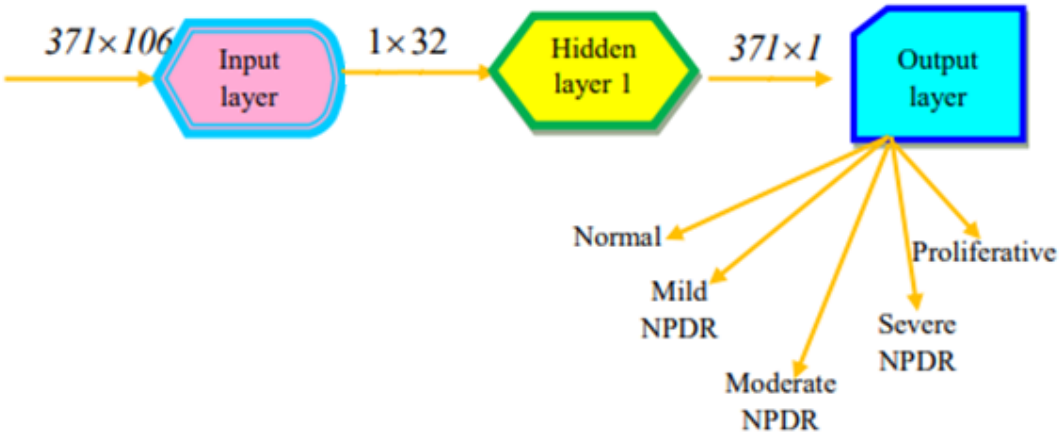

Figure 6.1: Architecture of RideNN

Architecture of RideNN

The output of the classifier is computed using a transfer function as

$$M_r = L^{L+2} * \left[\log \text{sig} \left(\sum_{v=1}^{I} B_r^v * L^v + L^{L+1} \right) \right] + L^{L+3} \tag{6.2}$$

where **logsig** represents the log-sigmoid transfer function that computes the output from the net input given; B_r^v is the input of the v^{th} neuron; L^v is the weight of the v^{th} neuron in the hidden layer and L^{L+2} is the output weight. Also, L^{L+1} and L^{L+3} are the biases. Hence, RideNN output is expressed as M_r and it classifies the DR.

Proposed Chronological Rider Sea Lion Optimization Approach to Train RideNN

The hybridization of the chronological concept with the SLnO and ROA is carried out for identifying optimal weights with minimal optimization issues and high convergence speed.

1) Initialization: Here, four rider groups and its locations are initialized randomly, which is indicated as $\mathcal{W}$. The group initialization is specified as,

$$R_q = \{R_q(f, g)\} \quad : \quad 1 \leq f \leq B, 1 \leq g \leq C \tag{6.3}$$

Where, B indicates amount of riders and it is similar to $\mathcal{W}$. C specifies dimension of coordinates, $R_q(f, g)$ represents location of f^{th} rider at time q. Moreover, the rider parameters, such as brake A_x, steering E_x, gear I_x, and accelerator N_x are initialized. Here, $R_q \in L^{L+2}, L^{L+1}, L^{L+3}$.

2) Fitness function calculation: The fitness measure is estimated to calculate optimal solution for DR classification. The fitness function with least value is considered as best solution, which is computed by,

$$\mathscr{G} = \frac{1}{\varphi} \sum_{r=1}^{\varphi} P_r^* - P_r \tag{6.4}$$

where, φ denotes total amount of samples, $\mathscr{G}$ symbolizes fitness measure, P_r^* specifies target output and P_r indicates classified output from *RideNN* classifier.

3) Leading Rider Detection:
After the estimation of the fitness function, the leading rider is determined based on the fitness value. The rider with the maximal fitness is selected as the leading rider. The position of the leading rider continuously changes in accordance with the fitness values.

$$\text{Leading Rider } (R_{\text{lead}}) = \arg\max_i \left(f(R_i)\right) \tag{6.5}$$

where $f(R_i)$ represents the fitness of the i^{th} rider.

4) Position Updation of Rider:
The position of each rider in a set is updated in order to identify the leading rider. The position updation depends on the rider's characteristics. The general position update

rule for rider R_i at iteration $t+1$ is given by:

$$R_i^{(t+1)} = R_i^{(t)} + \alpha \cdot \left(R_{\text{lead}}^{(t)} - R_i^{(t)} \right) + \beta \cdot \delta \tag{6.6}$$

where:

- $R_i^{(t)}$ is the current position of the i^{th} rider,
- $R_{\text{lead}}^{(t)}$ is the position of the leading rider,
- α is the learning factor or acceleration coefficient,
- $\beta \cdot \delta$ is a perturbation term modeling rider-specific behavior (e.g., randomness or directionality),
- δ is a small random vector drawn from a uniform or normal distribution.

a) Updation Process of Bypass Rider

The location updation in this set is specified on a random basis, because the bypass rider bypasses the common path deprived of subsequent leading riders. This is given by,

$$R_{g+1}^{B}(f,g) = \lambda \left[R_{\alpha}(g) \cdot \psi(g) + R_{\sigma}(g) \cdot (1 - \psi(g)) \right] \tag{6.7}$$

where,

- λ: random integer in the range $[0, 1]$,
- α: random number from 1 to B,
- σ: a number that selects a random value among 1 to B,
- ψ: a function returning a random integer in the range $[0, 1]$.

Therefore, every rider in the bypass group updates its location to become the winner.

b) Update Follower Position

The follower updates its location based on the leading *rider*, thus the rider rapidly reaches the target. Moreover, the follower mainly depends on a coordinate selector, so the location update is carried out for selected values in C.

$$R_{g+1}^{F}(f,g) = R^{O}(O,g) + \left[\cos\left(Z_{f,s}^{g} \right) * R^{O}(O,g) * a_f^g \right] \tag{6.8}$$

where,

- g symbolizes coordinate selector,
- R^O is the location of the leading rider,

- O implies the leading rider,
- $Z_{f,s}^{g}$ refers to the steering angle of the f^{th} rider in the g^{th} coordinate,
- a_{f}^{g} implies the distance traveled by the f^{th} rider.

The above equation is re-written in vector form:

$$\vec{R}(q+1) = \vec{R}(O) + \left[\cos(Z^{g}) * \vec{R}(O) * a^{g}\right] \tag{6.9}$$

$$\vec{R}(q+1) = \vec{R}(O)\left[1 + \cos(Z^{g}) \cdot a^{g}\right] \tag{6.10}$$

The circle updating location from SLnO algorithm is included, thus...

$$\vec{T}(q+1) = \left|\vec{E}(t) - \vec{T}(q)\right| \cos(2\pi u) + \vec{E}(t) \tag{6.11}$$

where, $\left|\vec{E}(t) - \vec{T}(q)\right|$ is the distance among search agent and the optimum solution, $||\cdot||$ refers to the absolute value, and u denotes a random integer from $[-1, 1]$. Here, $\vec{T}(q)$ indicates the position vector of the sea lion, and $\vec{E}(t)$ implies the position vector of the target prey.

Let us assume, $E(q) > T(q)$, $\vec{T}(q+1) = \vec{R}(q+1)$, $\vec{T}(q) = \vec{R}(q)$, and $\vec{E}(q) = \vec{R}(O)$,

$$\vec{T}(q+1) = \left(\vec{R}(O) - \vec{R}(q)\right) \cos(2\pi u) + \vec{R}(O) \tag{6.12}$$

$$\vec{T}(q+1) = \vec{R}(O)\cos(2\pi u) - \vec{R}(q)\cos(2\pi u) + \vec{R}(O) \tag{6.13}$$

$$\vec{T}(q+1) = \vec{R}(O)\left(\cos(2\pi u) + 1\right) - \vec{R}(q)\cos(2\pi u) \tag{6.14}$$

$$\vec{R}(O) = \frac{\vec{T}(q+1) + \vec{R}(q)\cos(2\pi u)}{\cos(2\pi u) + 1} \tag{6.15}$$

Substitute equation (6.15) in (6.10),

$$\vec{R}(q+1) = \frac{\vec{T}(q+1) + \vec{R}(q)\cos(2\pi u)}{1 + \cos(2\pi u)}\left[1 + \cos(Z^{g}) \cdot a^{q}\right] \tag{6.16}$$

$$\vec{R}(q+1) - \frac{\vec{T}(q+1)(1 + \cos(Z^{g}) \cdot a^{q})}{1 + \cos(2\pi u)} = \frac{\vec{R}(q)\cos(2\pi u)(1 + \cos(Z^{g}) \cdot a^{q})}{1 + \cos(2\pi u)} \tag{6.17}$$

$$\vec{R}(q+1)\left[\frac{1+\cos(2\pi u)-1-\cos(Z^g)\cdot a^q}{1+\cos(2\pi u)}\right]=\frac{\vec{R}(q)\cos(2\pi u)(1+\cos(Z^g)\cdot a^q)}{1+\cos(2\pi u)} \quad (6.18)$$

$$\vec{R}(q+1)=\frac{\vec{R}(q)\cos(2\pi u)(1+\cos(Z^g)\cdot a^q)}{\cos(2\pi u)-\cos(Z^g)\cdot a^q} \quad (6.19)$$

In addition, the chronological concept is incorporated, thus,

$$\vec{R}(q+1)=\frac{\vec{R}(q+1)+\vec{R}(q+1)}{2} \quad (6.20)$$

From equation (6.20) at iteration q,

$$\vec{R}(q)=\frac{\vec{R}(q-1)\cos(2\pi u)[1+\cos(Z^{g-1})\cdot a^{q-1}]}{\cos(2\pi u)-\cos(Z^{g-1})\cdot a^{q-1}} \quad (6.21)$$

Substitute equation (6.21) in (6.20),

$$\vec{R}(q+1)=\left(\frac{\vec{R}(q-1)\cos(2\pi u)[1+\cos(Z^{g-1})\cdot a^{q-1}]}{\cos(2\pi u)-\cos(Z^{g-1})\cdot a^{q-1}}\right)\left(\frac{\cos(2\pi u)[1+\cos(Z^g)\cdot a^q]}{\cos(2\pi u)-\cos(Z^g)\cdot a^q}\right) \quad (6.22)$$

$$\vec{R}(q+1)=\left(\frac{\vec{R}(q-1)\cos^2(2\pi u)[1+\cos(Z^{g-1})\cdot a^{q-1}][1+\cos(Z^g)\cdot a^q]}{(\cos(2\pi u)-\cos(Z^{g-1})\cdot a^{q-1})\,(\cos(2\pi u)-\cos(Z^g)\cdot a^q)}\right) \quad (6.23)$$

$$R(q+1)=\frac{1}{2}\left[\frac{\vec{R}(q)\cos(2\pi u)[1+\cos(Z^g)\cdot a^q]}{\cos(2\pi u)-\cos(Z^g)\cdot a^q}+\frac{\left(\frac{\vec{R}(q-1)\cos^2(2\pi u)[1+\cos(Z^{g-1})\cdot a^{q-1}][1+\cos(Z^g)\cdot a^q]}{\cos(2\pi u)-\cos(Z^{g-1})\cdot a^{q-1}}\right)}{\cos(2\pi u)-\cos(Z^g)\cdot a^q}\right] \quad (6.24)$$

$$R(q+1)=\frac{1}{2}\frac{\cos(2\pi u)[1+\cos(Z^g)\cdot a^q]}{\cos(2\pi u)-\cos(Z^g)\cdot a^q}\vec{R}(q)+\left[\frac{\vec{R}(q-1)\cos^2(2\pi u)[1+\cos(Z^{g-1})\cdot a^{q-1}]}{\cos(2\pi u)-\cos(Z^{g-1})\cdot a^{q-1}}\right] \quad (6.25)$$

The above expression is the updated equation by developed CRSLO approach for follower.

c) Updation process of overtaker

Here, the updation mainly depends on three essential elements, including relative success rate, direction indicator and coordinate selector. The expression of updation is denoted as,

$$R_{q+1}^{O}(f,g) = R_q(f,g) + \left|N_q^{\uparrow}(f)\right| * R^{O}(O,g) \tag{6.26}$$

where, $R_q(f,g)$ indicates location of f^{th} rider in g^{th} coordinate, $N_q^{\uparrow}(f)$ refers direction indicator of f^{th} rider at time q. Moreover, the distance indicator estimates the direction based on relative success rate,

$$N_q^{\uparrow}(f) = \left[\frac{2}{1-\log\left(A_q^g(f)\right)}\right] - 1 \tag{6.27}$$

where, $A_q^g(f)$ denotes relative success rate of f^{th} rider at period q, which is expressed as,

$$A_q^g(f) = \frac{x_g(f)}{\max\limits_{f-1} x_g(f)} \tag{6.28}$$

where, $x_g(f)$ represents success rate of f^{th} rider.

d) Modification of attacker location

The attacker location is same as the leader position, since attacker attempts to hold the location of leader. Moreover, location of attacker is updated by,

$$R_{q+1}^{A}(f,g) = R^{O}(O,g) + \cos(Z_{i_s}^q) * R^{O}(O,g) + a_f^q \tag{6.29}$$

where, $R^{O}(O,g)$ indicates location of leading rider, $Z_{i_s}^q$ is steering angle of f^{th} rider in g^{th} coordinate, and a_f^q implies distance travelled by f^{th} rider.

5) Fitness re-estimation: The maximum fitness value of the rider is contemplated as optimum solution, where equation (6.4) is utilized to identify the fitness measure.

6) Modification of rider parameter: The parameters is computed by considering the activity counter and is utilized to identify the optimal solution.

7) Termination: In *algorithm*, the pseudocode of CRSLO is depicted, which is used to identify optimal solution.

S. No.	Algorithm of proposed CRSLO model
1	Input: Rider's random position
2	Output: Optimal solution
3	Start
4	Set population
5	Set rider parameters
6	Determine fitness function by utilizing equation (6.4)
7	**while** $q < q_{sf}$
8	**for** $f = 1$ to B
9	Update location of bypass rider using equation (6.7)
10	Update location of follower by equation (6.25)
11	Update overtaker location based on equation (6.26)
12	Update location of attacker using equation (6.29)
13	Check feasibility of solution
14	Update rider parameters
15	Return optimal solution
16	$q = q + 1$
17	**end for**
18	**end while**

Thus, the incorporation of ROA with SLnO and chronological methods effectively classifies the DR types with less processing time. Here, the RideNN classifier categorizes the DR as mild NPDR, proliferative, normal, severe NPDR, and moderate NPDR.

6.4 Experimental Result

Table 6.1: Comparative discussion on DR Classification

Based on	Metrics	CNN	DCNN	SDL	Proposed
IDRID Dataset					
Training data percentage	Accuracy (%)	76.40	80.73	87.61	**92.18**
	Sensitivity (%)	78.95	82.35	90.10	**93.04**
	Specificity (%)	74.97	79.83	86.86	**91.54**
DDR Dataset					
Training data percentage	Accuracy (%)	75.18	79.52	86.21	**90.71**
	Sensitivity (%)	77.85	81.20	88.85	**91.55**
	Specificity (%)	73.78	78.56	85.47	**90.08**

The proposed model effectively segments the retina images using SeGAN model into types such as microaneurysm, Hemorrhages, hard exudates, and soft exudates, with a segmentation accuracy of 90.30%, which is more compared to other existing methods. Also,

the proposed model demonstrated improved performance in diabetic retinopathy (DR) classification, achieving an accuracy of 92.18%, a sensitivity of 93.04%, and a specificity of 91.54%. This robust performance is attributed to the integration of the Chronological Rider Sea Lion Optimization (CRSLO) with the Rider Neural Network (RideNN), which effectively identifies DR in its early stages. The segmentation model SeGAN faces challenges in capturing small and subtle lesions, such as microaneurysms and hemorrhages. Additionally, its performance is highly reliant on the quality and diversity of the training dataset. To address these limitations, a more diverse and well-annotated dataset is essential to improve the model's generalization capabilities. Here, performance is evaluated only on two datasets separately. To evaluate its adaptability and performance across different data sources, the proposed method is required to test on diverse datasets.

Chapter 7

Hybrid hunter-prey Ladybug beetle optimization enabled deep learning for diabetic retinopathy classification

7.1 Introduction

In this research, to overcome the limitation of proposed method 1, here used four diverse publicly available datasets along with real-time images. For early diagnosis of Diabetic Retinopathy, a novel optimized Hunter-Prey Ladybug Beetle Optimizer Deep Maxout Network (HPLBO_DMN) is proposed for DR classification. The input image is initially retrieved from the database and preprocessed. The Kalman filter and Region of Interest (RoI) extraction are used in the preprocessing stage to reduce noise in the input image. The preprocessed image is next subjected to a wavelet transform, where the Meyer wavelet is used to separate the image into subbands for additional processing. Thereafter, the lesion segmentation is done using K-Net, which is trained using HPLBO incorporated to avoid the limitation of SeGAN in identifying small and subtle lesions. The Automatic Artery/Vein (A/V) Classification Network (AVNet) classifies the arteries and veins from the wavelet-transformed image simultaneously. Additionally, the feature extraction phase obtains the output of the AVNet. Finally, DR classification is carried out with the help of a DMN. Here, DMN is trained using the proposed HPLBO, which is developed using the Hunter–Prey Optimizer (HPO) and the Ladybug Beetle Optimization (LBO). The introduced HPLBO_DMN attained higher performance of accuracy of 93.6%, sensitivity of 93.8%, and specificity of 92.9%. K-Net_HPLBO attained segmentation accuracy of 92.9% which shows performance improvement over SeGAN by 3%.

7.2 Proposed Methodology

Image acquisition

Let us assume input image from a database $\mathcal{B}$ with e number of images, which is detailed as,

$$\mathcal{B} = \{B_1, B_2, \ldots, B_u, \ldots, B_e\} \tag{7.1}$$

Here, the database implied as $\mathcal{B}$, total number of images in database is indicated as e and B_u represents u^{th} image in a database.

Image pre-processing step

The input image is taken from a dataset and pre-processed with a Kalman filter to remove noise and outliers. As a consequence, the quality of input image is significantly enhanced.

Kalman filter for image pre-processing

The most popular Bayes filter variation is the Kalman filter [32]. The fundamental benefit of Kalman filters is their computational effectiveness. It is only possible to use Kalman filters for location tracking with either precise sensors or sensors with high update rates when the state uncertainty is not too high. The best estimators are Kalman filters when the system dynamics and observation model are both linear functions of the state, and the starting uncertainty is assumed to be Gaussian. Typically, extended Kalman filters are used to linearize systems because most schemes are not exactly linear. These filters employ expansions from the first-order Taylor series. Here, Kalman filter for input image is specified as,

$$\psi(f_{\mathbf{w}}) \approx \mathcal{G}(f_{\mathbf{w}}; \lambda_{\mathbf{w}}, \Sigma_{\mathbf{w}}) \tag{7.2}$$

Wherein, mean distribution is represented as $\lambda_{\mathbf{w}}$, $\Sigma_{\mathbf{w}}$ stands for estimate uncertainty; the greater the covariance, the greater the spread of the distribution. Moreover, $\mathcal{G}(f_{\mathbf{w}}; \lambda_{\mathbf{w}}, \Sigma_{\mathbf{w}})$ specifies the probability of $f_{\mathbf{w}}$, given a Gaussian with mean $\lambda_{\mathbf{w}}$, and covariance $\Sigma_{\mathbf{w}}$.

ROI extraction

By using the ROI extraction [29] technique to extract a specified area of an image, the computational challenges are significantly reduced. The ROI is calculated by taking into consideration the pixel intensity values of pre-processed image. The ROI extraction method is employed to identify critical regions in a noisy image. To enable an effective classification approach, the dispersed component is here divided from non-responsive area. The boundary for the introduced section of an input image is defined by ROI,

or pixel intensity value. The pre-processed images have external elements removed to remove external misrepresentations. The result of the ROI extraction is given as ζ_z.

Wavelet transform

Here, outcome of ROI extraction ζ_z is taken as input for wavelet transform. The next stage of lesion segmentation employs wavelet transform. The wavelet transform often refers to the multi-resolution decomposition procedure when discussing image expansion in a collection of wavelet basis functions. Meyer wavelet [33] has limitless support and is orthogonal and differentiable. According to the frequency domain definition it is stated in equation (7.3),

$$\varepsilon(I) = \begin{cases} (2\pi)^{-\frac{1}{2}} \cdot \sin\left(\frac{\pi}{2} h\left(\frac{3}{2\pi}|I| - 1\right)\right) & \text{if } \frac{2\pi}{3} \leq |I| \leq \frac{4\pi}{3} \\ (2\pi)^{-\frac{1}{2}} \cdot \cos\left(\frac{\pi}{2} h\left(\frac{3}{2\pi}|I| - 1\right)\right) & \text{if } \frac{4\pi}{3} \leq |I| \leq \frac{8\pi}{3} \\ 0 & \text{otherwise} \end{cases} \tag{7.3}$$

This yields wavelets and scaling functions with limitless support, but they decay faster than sine wavelets.

$$h(w) = \begin{cases} 0, & \text{if } w < 0 \\ w, & \text{if } 0 < w < 1 \\ 1, & \text{if } w > 1 \end{cases} \tag{7.4}$$

Meyer wavelet is used to further divide the images into low- and high-frequency segments. This section contains high-frequency data, while the next section further divides low-frequency data into high-frequency and low-frequency components. Once the first level decomposition is complete, the image is divided into four *subbands*, as follows.

$$\kappa(\zeta_z) = \{HH, HL, LH, LL\} \tag{7.5}$$

Here, wavelet transform is indicated as κ; high-high, high-low, low-high, low-low wavelet coefficients are indicated as HH, HL, LH, LL. Despite, the wavelet coefficient of HL has more accurate data concerning image for lesion segmentation procedure. The outcome attained from wavelet transform is signified as α_z.

Lesion segmentation using K-Net

In this process, wavelet transform α_z output is used as the starting point for lesion segmentation. After that, the lesion area is segmented during lesion segmentation based on K-Net [34]. The K-Net architecture constantly separates instances from semantic categories using a set of learnable kernels, each of which is in charge of producing a

mask for either a possible instance or a stuff class. The maximum group size for a segmentation task is set to K since it is usual to assume that the number of groups in an image is finite. Definitely, input feature map $b \in \mathbb{S}^{R\times k\times m\times p}$ of R images, formed by deep neural network, but only required K kernels $t \in \mathbb{S}^{K\times k}$ to execute convolution with b to attain the equivalent segmentation calculation $z \in \mathbb{S}^{R\times K\times m\times p}$ as,

$$z = \omega(t * b) \tag{7.6}$$

Here, number of width of feature map is signified as p, number of height is designated as m and number of channels is signified as k correspondingly. Moreover, activation function ω can be soft max function and it is necessary to allocate every pixel to only one of the kernels. In this instance, the K-Net classifies the lesion into four categories namely hard exudates, microaneurysms, hemorrhages, and soft exudates. The output of K-Net is denoted as J_z.

Training algorithm using proposed HPLBO_DMN

The K-Net's approach for choosing its best region is enhanced by applying the HPLBO technique to improve its trainable factors. Merging the HPO [35] and LBO [36] algorithms result in the HPLBO algorithm. The population-based optimization algorithm HPO is also influenced by the behavior of predatory animals like lions, leopards, and wolves as well as prey species like deer and gazelles. The HPO algorithm provides an excellent balance between exploration and exploitation stages, making it capable of handling both unimodal and multimodal situations. Additionally, LBO draws inspiration from Lady Bird and natural wintertime behavior of seeking out a warm location. LBO experienced the lowest power losses within the grids. LBO was also used to forecast the COVID-19 model as a real-world problem. The findings showed that LBO performed better than other algorithms in meta-heuristic frameworks. Such a hybridized approach is designated as HPLBO. Below is a description of the HPLBO algorithm's steps.

Step 1): Population initialization

For each member of the initial population in the search space, equation (7.7) creates a random position.

$$E_w = rand(1, F) * (\varphi - \eta) + \eta \tag{7.7}$$

Wherein, hunter prey or position is represented as E_w, the problem variables lower boundary η is their lowest value, while their upper boundary φ is their highest value and the dimensions or number of variables of the problem is designated as F.

Step 2): Objective function evaluation

When hunter prey position is determined, the fitness parameter for the points is also determined. Here, optimal region matches response, therefore fitness is designed using

K-Net output and is given as,

$$C = \frac{1}{D}\sum_{n=1}^{D}(J_n^t - J_n)^2 \tag{7.8}$$

Wherein, mean square error is represented as C, K-Net output is indicated as J_n, expected output is signified as J_n^t, and total number of training samples is represented as D.

Step 3): The search mechanism

Exploration and exploitation are the two processes that typically make up search mechanism. The algorithm's penchant for exceptionally unpredictable behavior, known as exploration, causes solutions to vary significantly. A deeper investigation of the search space and identification of its promising regions are prompted by the major changes in solutions. For the hunter search mechanism,

$$E_{w_k}(g+1) = E_{w_k}(g) + 0.5\left[\left(2HNN_{pos(g)} - E_{w_k}(g)\right) + 2(1-H)Y_{th(g)} - E_{w_k}(g)\right] \tag{7.9}$$

Here, adaptive parameter is represented as Y, current position is signified as $E(g)$, next hunter position is denoted as $E(g+1)$, prey position is indicated as N_{pos}, mean of position is denoted as $\hat{h}$.

$$N = \vec{P}_1 < H; \vec{\xi} = (N == 0) \tag{7.10}$$

$$Y = P_2 \otimes \vec{\xi} + P_3 \otimes (\sim \vec{\xi}) \tag{7.11}$$

Here, random vectors are signified as $\vec{P}_1$ and $\vec{P}_3$, index number of the vector is denoted as $\vec{\xi}$, structure of parameter is represented as Y, and random number in the range $[0, 1]$ is denoted as P_2.

Step 4): Balance parameter and distance

Exploration and exploitation are balanced by a parameter called H, whose value falls over the course of iterations from 1 to 0.02.

$$H = 1 - \frac{iter}{\left(\frac{0.98}{\text{max iter}}\right)} \tag{7.12}$$

Here, current iteration value is represented as *iter* and maximum number of iterations is denoted as *max iter*. Moreover, an average of all positions is expressed in equation (7.13).

$$\hat{h} = \frac{1}{T}\sum_{w=1}^{T}\vec{E}_w \tag{7.13}$$

Equation (14), calculates the distance using the Euclidean principle.

$$V_{eval(w)} = \left(\sum_{w=1}^{T} \left(\vec{E}_w - \hat{h} \right)^2 \right)^{\frac{1}{2}} \tag{7.14}$$

The search agent with greatest deviation from mean of positions is referred to as the prey $\overline{N}_{pos}$ in equation (15).

$$\overline{N}_{pos} = \vec{E}_w \mid W = \text{index of } \max(\text{end}) \text{ sort}(V_{eval}) \tag{7.15}$$

$$\beta_{best} = \text{round}(H \times \tau) \tag{7.16}$$

$$\overline{N}_{pos} = \vec{E}_w \mid W = \text{sorted}(V_{eval} < \beta_{best}) \tag{7.17}$$

Here, τ represents the number of search engines.

$$E_{w,k}(g+1) = L_{\text{pos}(k)} + HT \cos(2\pi P_k) \times (L_{\text{pos}(k)} - E_{w,k}(g)) \tag{7.18}$$

$$E_{w,k}(g+1) = L_{\text{pos}(k)} + HT \cos(2\pi P_k) \times L_{\text{pos}(k)} - HT \cos(2\pi P_k) \times E_{w,k}(g) \tag{7.19}$$

$$E_{w,k}(g+1) = L_{\text{pos}(k)} \left[1 + HT \cos(2\pi P_k)\right] - HT \cos(2\pi P_k) E_{w,k}(g) \tag{7.20}$$

The efficacy of HPLBO algorithm is improved with the usage of LBO [36]. The equation for LBO is as follows:

$$Q_j(C+1) = Q_j(C) + rand \times (Q_j(C) - Q_j(C)) + rand \times (Q_j(C) - Q_{L_r}(C)) + rand \times |M_j| \frac{C}{Z(C)} Q_j(C) \tag{7.21}$$

Let us consider, $Q_j(C+1)$ as $E_{w,k}(g+1)$, $Q_j(C)$ as $E_{w,k}(g)$, $Q_j(C)$ as $E_c(g)$, and $Q_j(C)$ as $E_{c_L}(g)$.

The equation becomes,

$$E_{w,k}(g+1) = E_{w,k}(g) + rand(E_c(g) - E_{w,k}(g)) + rand(E_c(g) - E_{c_L}(g)) + rand\, |M_j| \frac{C}{Z(C)} Q_j(C) \tag{7.22}$$

$$E_{w,k}(g+1) = E_{w,k}(g)\left(1 - rand + rand\,|M_j|\,\frac{C}{Z(C)}\right) + rand \times E_c(g) + rand(E_c(g) - E_{c_L}(g)) \tag{7.23}$$

$$E_{w,k}(g) = \frac{E_{w,k}(g+1) - rand \times E_c(g) + rand(E_{c_L}(g) - E_c(g))}{1 - rand + rand\,|M_j|\,\frac{C}{Z(C)}} \tag{7.24}$$

Substitute equation (7.24) in equation (7.20), and the equation becomes,

$$E_{w,k}(g+1) = L_{\text{pos}(k)}\left[1 + HT\cos(2\pi P_k)\right] - HT\cos(2\pi P_k)\left[\frac{E_{w,k}(g+1) - \text{rand} \times E_c(g) + \text{rand}\left(E_{c_L}(g) - E_c(g)\right)}{1 - \text{rand} + \text{rand}|M_j|\,\frac{C}{Z(C)}}\right] \tag{7.25}$$

$$\begin{aligned} &E_{w,k}(g+1) + \frac{HT\cos(2\pi P_k)E_{w,k}(g+1) - (HT\cos(2\pi P_k))\left[\text{rand}\left(E_{c_L}(g) - E_c(g)\right) - \text{rand} \times E_c(g)\right]}{1 - \text{rand} + \text{rand}\,|M_j|\,\frac{C}{Z(C)}} \\ &\quad = L_{\text{pos}(k)}\left[1 + HT\cos(2\pi P_k)\right] - \left\{1 - \text{rand} + \text{rand}\,|M_j|\,\frac{C}{Z(C)}\right\} \end{aligned} \tag{7.26}$$

$$\begin{aligned} &\left[\frac{1 - \text{rand} + \text{rand}\,|M_j|\,\frac{C}{Z(C)}}{HT\cos(2\pi P_k)}\right] E_{w,k}(g+1) \\ &\quad = \frac{L_{\text{pos}(k)}\left[1 + HT\cos(2\pi P_k)\right] - \left\{1 - \text{rand} + \text{rand}\,|M_j|\,\frac{C}{Z(C)}\right\}}{1 - \text{rand} + \text{rand}\,|M_j|\,\frac{C}{Z(C)}} \\ &\qquad - \frac{HT\cos(2\pi P_k)\left[\text{rand}\left(E_{c_L}(g) - E_c(g)\right) - \text{rand} \cdot E_c(g)\right]}{1 - \text{rand} + \text{rand}\,|M_j|\,\frac{C}{Z(C)}} \end{aligned} \tag{7.27}$$

$$\begin{aligned} E_{w,k}(g+1) = &\frac{L_{\text{pos}(k)}\left[1 + HT\cos(2\pi P_k)\right] - \left\{1 - \text{rand} + \text{rand}\,|M_j|\,\frac{C}{Z(C)}\right\}}{1 - \text{rand} + \text{rand}\,|M_j|\,\frac{C}{Z(C)} + HT\cos(2\pi P_k)} \\ &- \frac{HT\cos(2\pi P_k)\left[\text{rand}\left(E_{c_L}(g) - E_c(g)\right) - \text{rand} \cdot E_c(g)\right]}{1 - \text{rand} + \text{rand}\,|M_j|\,\frac{C}{Z(C)} + HT\cos(2\pi P_k)} \end{aligned} \tag{7.28}$$

Where, optimum global position is denoted as $L_{\text{opt}(g)}$, $\tau(g)$ represents the g^{th} ladybug from the distance between $(0, 1)$, and g indicates g^{th} iteration.

Step 5): Re-evaluate the fitness function

The fitness value for each individual is calculated using fitness function in equation (7.8), with the value with the highest fitness serving as the ideal result.

Step 6): Termination

Each of the above-mentioned procedures can be carried out a maximum number of times before the best option is identified. Algorithm displays pseudocode for HPLBO algorithm.

Algorithmic steps of HPLBO technique

1.	Start
2.	Set population randomly with equation (7.7)
3.	Input HPO parameters, $\max_{iter}$
4.	Determine fitness function with equation (7.8)
5.	Estimate L_{pos}
6.	Update H with equation (7.12)
7.	Evaluate Y with equation (7.10) and equation (7.11)
8.	$P_s < \beta$
9.	If the condition is satisfied
10.	Evaluate N_{pos} with equation (7.13)
11.	Update position with equation (7.28)
12.	If the condition is not satisfied
13.	Evaluate fitness function with equation (7.8)
14.	Evaluate L_{pos}
15.	If the termination condition met
16.	If the condition is satisfied
17.	Output the best solution
18.	End

As a result, using HPO and LBO approaches together efficiently distinguishes the different DR kinds. The DMN classifier classifies the DR in this case as mild NPDR, normal, severe NPDR, proliferative, and moderate NPDR.

Artery and vein segmentation using AVNet

Afterward, wavelet transformed outcome is considered as input for artery and vein segmentation process. Here, AVNet [31] is applied to the image to segment the artery and vein. To reduce feature maps and extract high-level semantic features, an encoder module is used. Using a decoder module, the precise spatial information can be recovered.

i) Encoder: The AVNet encoder uses GoogLeNet, which was trained on ImageNet beforehand. GoogLeNet has three inception modules, five pooling, three convolutional, and one fully connected layer. The inception module has multi-scale convolution operations that can capture the variation in artery and vein diameter.

ii) Decoder: The outputs of Inceptions 4 and 5 are used by the decoder to produce two side outputs, each with three channels. Since Inception 4 features keep more detailed data than Inception 5 features, skip-connection aims to transfer features from Inception 4 to Inception 5. Two steps make up this process. After the vertical features of inception4 have been applied, a convolution operation using kernel size is carried out. Second, vertical features of Inception 5 are concatenated to the obtained feature map to create a feature map with 17 channels.

iii) Loss function: The AVNet has three losses that have been optimized, including one output loss and two deep supervision losses. These three losses are added together to form the objective function. To address the issue of class imbalance, each loss is a cross-entropy loss function that is class-balanced, with distinct categories assigned varying weights. The following is a formulation for the class-balanced cross-entropy loss function.

$$G(i,a) = -\sum_{\rho=1}^{\infty-1}\sum_{\varnothing=1}^{\varnothing} X_{\circ}(\chi_{\rho})\zeta_{\circ}(\chi_{\rho})\log(i_{\circ}(\chi_{\rho})) \tag{7.29}$$

Here, height is represented as O, weight is designated as X, number of categories are showed as ϖ, predicted probability is indicated as $i_{\omega}(\chi_{\rho})$, and weight of each category is represented as $X_{\omega}(\chi_{\rho})$. Additionally, the weight settings for all three losses are the same. According to equation (7.30), the indicator function $\zeta_{\omega}(\chi_{\rho})$ indicates if the expected outcome is consistent with its real label.

$$a_{\omega}(\chi_{\rho}) = \begin{cases} 0, & \gamma(\chi_{\rho}) \neq \omega \\ 1, & \gamma(\chi_{\rho}) = \omega \end{cases} \tag{7.30}$$

Here, $\gamma(\chi_{\rho})$ indicates true label of pixel χ_{ρ}. The output of AVNet is denoted as $\hat{h}_{\tau}$.

Feature Extraction

Then, from the output of the artery and vein classification procedure, key characteristics are extracted. Artery-to-vein ratio, colour histogram, eccentricity, area, kurtosis, perimeter, LVP, entropy, and LGBP are some of the notable features that are extracted in this case.

DR Classification Using HPLBO_DMN

The extracted feature vector is considered as an input to the DMN model. In addition, the DMN is tuned using the HPLBO algorithm, which was created by combining the frameworks of LBO and HPO.

1. Architecture of DMN

In this section, the feature vector is fed as an input to the DMN. The DMN [37] is used for the DR classification process because it successfully reduces overfitting concerns and requires less processing time. Unlike only a feed-forward structure, like deep CNN or multilayer perceptron, the Maxout system employs a more recent class of activation function called a Maxout unit. Multiple layers make up DMN, which uses the Maxout function to produce hidden activations.

a) Rectified Linear Unit Layer (ReLU): The **ReLU** is first utilized in Restricted Boltzmann Machines (RBM), which are defined as,

$$p_r = \begin{cases} W_r, & \text{if } W_r \geq 0 \\ 0, & \text{if } W_r < 0 \end{cases} \tag{7.31}$$

Here, input to neuron is represented as W_r, and output is denoted as p_r. Moreover, **ReLU** has the appealing properties of sparse neural node activations.

b) Maxout Layer: The **Maxout** is a more generic implementation of **ReLU** that obtains the maximum function for $Y(Y = 2)$ trainable linear functions. Here Z is either state vector or the raw input vector of the hidden layer, for the known input $Z \in \mathbb{C}^T$. The following provides an illustration of a **Maxout** unit output:

$$\Gamma_r(Z) = \max_{n \in [1,Y]} S_n \tag{7.32}$$

Maximum across Y feature maps is what the CNN defines as **Maxout** unit activation. Although the maximum value over comparable input is obtained by the **Maxout** unit, which is equal to the spatial Max pooling typically used in CNN, spatial Max pooling is related with Y multiple input.

c) DMN: It is regarded as a variety of trainable activation function with multiple layer configurations. In the case of known input $Z \in \mathbb{C}^T$, where Z is either the state vector or the raw input vector of the hidden layer, an activation of the hidden unit is modelled by,

$$\delta^1_{r,u} = \max_{n\in[1,Y]} Z^T W^-_{rn} + I_{rn} \tag{7.33}$$

$$\delta^2_{r,u} = \max_{n\in[1,Y]} \delta^1_{r,u} W^-_{rn} + I_{rn} \tag{7.34}$$

$$\delta^3_{r,u} = \max_{n\in[1,Y]} \delta^2_{r,u} W^-_{rn} + I_{rn} \tag{7.35}$$

$$\delta^4_{r,u} = \max_{n\in[1,Y]} \delta^3_{r,u} W^-_{rn} + I_{rn} \tag{7.36}$$

$$\Gamma_r = \max_{n\in[1,Y]} \delta^4_{r,u} \tag{7.37}$$

Wherein, number of units in the ϕ^o layer signified as Y_ϕ while whole layers exist in DMN is signified as n. When compared to non-convex activation functions, DMN activation is significantly more powerful to approximate any given recurrent activation function. DMN approximates the traditional non-linear activation functions, such as the rectified linear and absolute value rectifier, where Y is not less than 2. DMN may approximate any activation function by increasing a parameter Y, however effective feature extraction still necessitates a highly complex non-linear function. Π provides a detected output, and it categorizes DR as severe NPDR, normal, mild NPDR, moderate NPDR, and proliferative. Figure 7.1 shows the diagrammatic structure of DMN network.

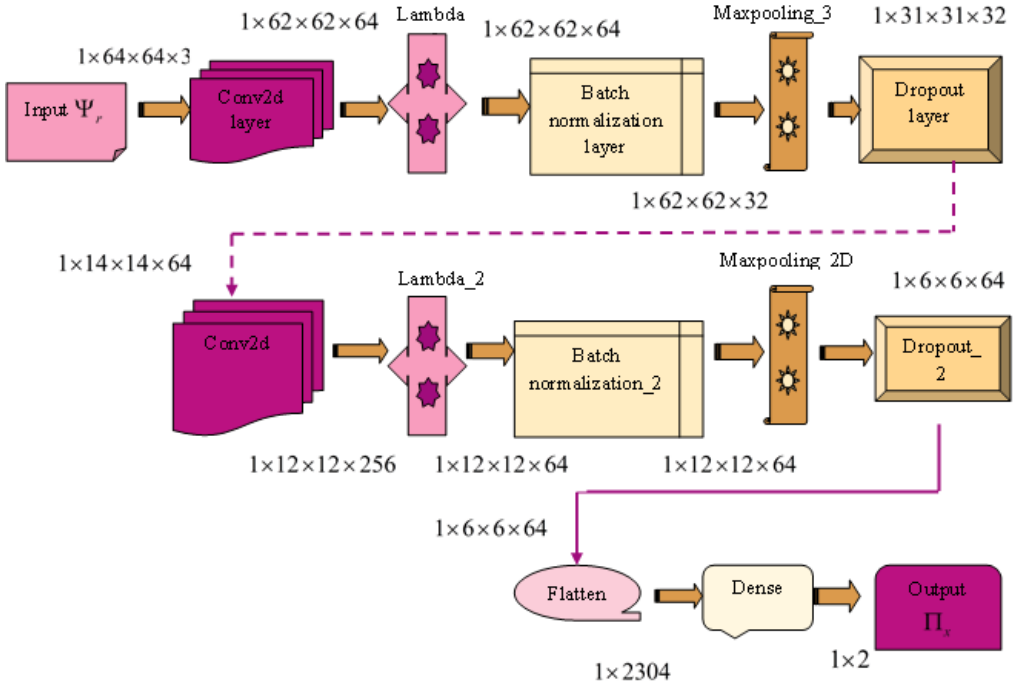

Figure 7.1: Architectural diagram of DMN

Training of DMN using HPLBO for DR classification

The created HPLBO, a hybrid optimization that includes the benefits of both HPO and LBO, is used to enhance the DMN weights and parameters. The HPLBO training algorithm is explained systematically earlier. The fitness parameter used is given by,

$$C = \frac{1}{D}\sum_{x=1}^{D}\left(J_x^* - \Pi_x\right)^2 \tag{7.38}$$

Wherein, an expected result is denoted as J_x^* and output obtained from the DMN is indicated as Π_x respectively. Moreover, the total number of samples is signified as D.

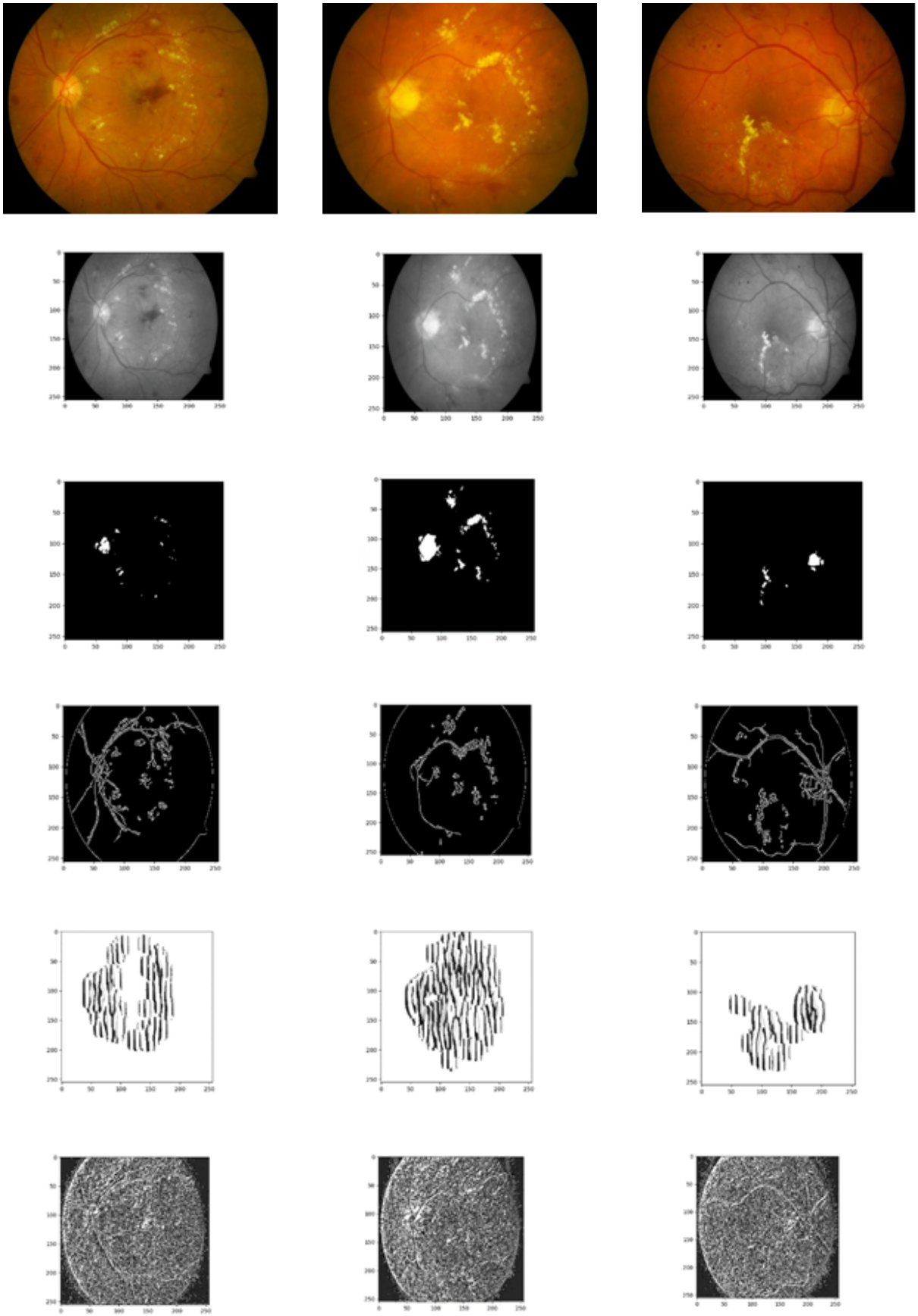

Figure 7.2: Experimental results of developed HPLBO_DMN a) Input image for 1, 2 and 3 datasets, b) Pre-processed image for 1, 2 and 3, c) lesion segmentation image for 1, 2, and 3,d) artery to vein segmented image 1, 2 and 3, e) LGBP image for dataset 1,2 and 3, f) LVP image for dataset 1,2 and

7.3 Experimental Analysis

The DR classification approaches, including CNN [27], Deep CNN[38], SDL [23], and CRSLO-based RideNN are considered for comparing the performance of devised HPLBO_DMN.

Comparative Analysis

Table 7.2: Comparative discussion on DR Classification

Based on	Metrics	CNN	DCNN	SDL	CRSLO RideNN	HPLBO DMN
Training set	Accuracy (%)	80.5	82.0	84.2	87.2	**93.6**
	Sensitivity (%)	78.4	83.2	85.5	88.5	**93.8**
	Specificity (%)	80.0	83.8	85.2	86.7	**92.9**
K value =9	Accuracy (%)	78.1	83.0	84.1	88.4	**92.6**
	Sensitivity (%)	78.3	82.4	84.9	88.1	**92.6**
	Specificity (%)	80.6	82.3	84.2	87.4	**93.2**

Comparative Discussion for Segmentation Accuracy

Table 7.3: Comparative discussion on DR segmentation accuracy

Based on	WT RBFNN [46]	LD DCNN [15]	Graphcut NN [83]	SeGAN [51]	HPLBO K-Net
Training SetAccuracy (%)	79.8	82.5	84.5	88.6	**91.8**
K-Fold Accuracy (%)	79.3	82.0	85.9	88.3	**92.9**

Conclusion

In this work, an effective DR classification framework named the HPLBO_DMN method is introduced. The model successfully combined Hunter-Prey Optimization (HPO) and Lady Beetle Optimization (LBO) techniques to enhance the accuracy and robustness of DR detection. The integration of these complementary metaheuristic algorithms enabled the model to achieve improved performance, with an accuracy of 93.6%, sensitivity of 93.8%, specificity of 92.9%, and segmentation accuracy of 92.9%. The SeGAN model achieves a segmentation accuracy of 88.6% on diverse datasets. The HPLBO K-Net demonstrates a 4.86% improvement in segmentation accuracy over SeGAN as shown in Table 7.3, indicating its superior performance in segmentation tasks. For future work, this research aims to incorporate more precise feature extraction methods and advanced optimization techniques to further refine the classification process and improve the early

detection of DR. This would enhance the model's effectiveness in diverse clinical settings, ultimately contributing to better patient outcomes.

Chapter 8

CAViaR Dung Beetle Optimization Enabled Deep Learning for Diabetic Retinopathy Classification

8.1 Introduction

CAViaR Dung Beetle Optimization-DenseNet (CDBO_DenseNet) is proposed in this research for DR classification. At first, an image is obtained from the mentioned 7 diverse datasets along with low and high-resolution real-time images. Images is given to a preprocessing module to enhance the image employing a mean Filter. Afterward, lesion segmentation is performed using Deep Joint Segmentation. Similarly, the process of classifying artery and veins is performed using AVNet. Subsequently, the feature extraction process extracts some essential features. Finally, DR is classified by DenseNet, which is tuned with CDBO. However, CDBO is designed by integrating DBO with the CAViaR technique.

8.2 Proposed Methodology

Fundus images contain diverse lesions. Each one has complex shapes and textures. Also, irregular boundaries and varying lesion sizes. Hence, the traditional method of segmentation is unable to handle multiple classes simultaneously. By considering the limitations of earlier proposed methods, here utilized the deep joint segmentation technique specifically for the segmentation of four dissimilar types of lesions, including Hard Exudates, Hemorrhages, Micro aneurysms, and Soft Exudates. fundus images. DenseNet is used for the classification of diabetic retinopathy as it can generalize well on intricate patterns and subtle features in retina images. To address the overfitting challenges of DenseNet, this research proposes a novel optimization framework, CAViaR Dung Beetle Optimization-

DenseNet (CDBO_DenseNet), specifically designed for DR classification. Compare all the proposed methods on the diverse datasets mentioned in Chapter 10 to evaluate their performance, robustness, and generalization capabilities across varying data sources and image resolutions.

Image Acquisition

Considering a database I having m number of images, the input image set is formulated as:

$$I = \{I_1, I_2, \ldots, I_n, \ldots, I_m\} \tag{8.1}$$

Here, the database is represented as I, where m denotes the total number of images, and I_m denotes the m^{th} image in the dataset.

Image Pre-processing

The input image I_n, acquired from the dataset, is pre-processed to remove noise and artifacts present in the image. As a result, the quality of the input image is considerably improved.

Here, image pre-processing is performed using a mean filter. The mean filter [39] enhances the value of each pixel in the image using the mean of grayscale values in its neighborhood.

$$B(y, z) = \frac{1}{c \times d} \sum_{x,y \in S_{yz}} J(x, y) \tag{8.2}$$

Here, S_{yz} denotes the neighborhood sub-image window of size $c \times d$, centered at the point (y, z), and the pre-processed image is denoted by B_n.

Lesion segmentation

Here, the output of the mean filter fed to lesion segmentation and the segmentation process is performed employing Deep joint segmentation [40] where lesion region is segmented. It contains three units, like joining stage, region fusion stage, together with segmentation point production stage. Firstly, grids are separated from the input image then subsequently; joining phase merges the pixel depending on mean and threshold calculation. Thereafter, region fusion is done considering the similarity of the region as well as bi-constraints. In segmentation point production stage, fine segments are projected considering the distance value between deep point as well as segmentation point together with calculated mean square error (MSE) value.

Formation of Grids

The input image is partitioned into multiple grids. Each grid has a dimension of 2×2. The partitioned grids of the preprocessed image B_n are represented as:

$$E = \{E_1, E_2, \ldots, E_i, \ldots, E_j\} \tag{8.3}$$

Here, E_i represents the i^{th} grid, and j is the total number of grids.

Joining Phase

Following grid creation, intra-grid pixels are grouped based on a threshold and mean value. By considering the average value of pixels in a grid, a mean rate is computed and processed using a fixed threshold (set to 1). The average value is calculated as:

$$E_i = \frac{\sum_{i=1}^{N} P_i}{N} \tag{8.4}$$

where P_i denotes the pixel value corresponding to grid E_i, and N is the number of pixels in the grid E_i.

The equation for pixel joining is given as:

$$E_i = \frac{\sum_{i=1}^{N} P_i}{N} \pm t \tag{8.5}$$

where t denotes the threshold.

Region Fusion Phase

This phase is formulated based on the allocated grid. Subsequently, regions are fused by analyzing the similarity in pixel intensity and applying bi-constraints to identify resemblance across regions. The formulation is expressed as:

$$E_i = \frac{\sum_{o=0}^{w} P_o^r}{R} \tag{8.6}$$

Here, P_o^r represents the joined pixel in a particular grid E_i, and R denotes the total number of joined pixels in E_n.

The fused grid is mapped as:

$$R = \{R_1, R_2, \ldots, R_t\} \tag{8.7}$$

where t represents the number of mapped points in the image.

Identification of Deep Points

After that, by considering the missed pixels from the joining phase and the mapped points, deep points are predicted. Pixels left over in the image after the joining phase — or those not falling within the threshold boundary — are termed as Q. These pixels are represented as:

$$F = \{Q_H\}, \quad 1 < H \leq D \tag{8.8}$$

Here, Q_H is the pixel missed during the joining phase, and D denotes the total number of missed and mapped pixels.

Optimal Segment searching

Lastly, an optimal segment is forecasted considering deep points using an iterative strategy, and it is represented as,

$$X^{Dis} = \sqrt{\sum_{i=1}^{20} (Z_i - D_{\text{point}_i})^2} \tag{8.9}$$

This procedure is repeated, and the best-segmented points are chosen employing MSE, which analyzes the similarity among the two grids. The output from the lesion segmentation is L_n.

Artery and vein segmentation using AVNet

Consequently, the preprocessed image is fed to the Artery and vein segmentation module. Now, AVNet [16] is employed to segment the artery and vein, and it follows an encoder-decoder architecture.

Encoder

This encoder unit is used in obtaining large levels of semantic features as well as to decrease the feature maps. GoogleNet [38] is employed in the encoder module in AVNet. It has three beginning stages, three convolutions, one fully connected layer, and five pooling layers.

Decoder

This unit is used to combine the features from various level as well as to get extra refined segmentation outcomes and it obtain output of inception4, and inception5 as input in addition it generates two side output every one having three channels.

Category-attention weighted fusion (CWF)

To improve the classification ability of *AVNet* the CWF is merged, and the decoder produces two side outputs $\{A^{(1)}, A^{(2)}\}$, where each of them has K channels. Different from easy concatenation, CWF performs convolution in every group individually, by avoiding communication between features from various types:

$$A^f = \{A_1^{(1)}, A_1^{(2)}, A_2^{(1)}, A_2^{(2)}, ..., A_k^{(1)}, A_k^{(2)}\} \tag{8.10}$$

Here, K-grouped 1×1 convolution is performed on A^f to obtain final segmentation output.

Loss function

Three losses optimized in this *AVNet* are two deep supervision losses and one output loss. Cross-entropy loss function formulated is shown below as,

$$K(p, y) = -\sum_{i=1}^{H \times W} \sum_{c=1}^{K} w_c(x_i)\, y_c(x_i) \log(p_c(x_i)), \tag{8.11}$$

Here K denotes the number of categories, $p_c(x_i)$ denotes the predicted probability and $w_c(x_i)$ is the weight of each category. Hence, allotment of vessel/non-vessel pixels is greatly inclined as well as arteries are simply misclassified into veins. Additionally, the above losses employ similar setting of weight. $y_c(x_i)$ is considered as a display utility to explain if detected outcome is steady by their accurate label.

$$y_c(x_i) = \begin{cases} 1, & l(x_i) = c \\ 0, & l(x_i) \neq c \end{cases} \tag{8.12}$$

$l(x_i)$ indicates the true label of pixel x_i. Outcome of *AVNet* is expressed as W_n.

Feature extraction

It extracts some important features in segmented image. Here from output of *AVNet* W_n, the important features extracted such as Artery to vein ratio, Area, Entropy, eccentricity, area, kurtosis, perimeter, LVP, entropy, and LGBP.

DR classification using DenseNet

DR classification is carried out using **DenseNet**[41] considering the output from the feature extraction F_n. **DenseNet** is trained using the proposed CDBO, which is the integration of Dung DBO with **CAViaR** technique.

Architecture of DenseNet This module captures the spatial correlation among the regions. It is very simple and narrow and it improved the flow of information and gradient throughout the region.

(a) ResNet The skip-connections are summed, allowing identity mapping and avoiding nonlinear transformations. The transformation is represented as:

$$x_l = H_l(x_{l-1}) + x_{l-1} \tag{8.13}$$

(b) Dense Connectivity In DenseNet, the entire preceding layers' feature maps are concatenated as inputs to the l-th layer. The transformation is given by:

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \tag{8.14}$$

(c) Composite Function The function $H_l(\cdot)$ is composed of three consecutive operations: a 3×3 convolution, followed by batch normalization (BN), and a ReLU activation.

(d) Pooling Layers Both pooling and convolution operations are achieved by the transition layers. Each transition layer includes a 1×1 convolution followed by a 2×2 average pooling layer. Additionally, batch normalization (BN) is applied after the convolution.

(e) Growth Rate The growth rate controls how much new information each layer contributes. It ensures that global condition information, once introduced, can propagate throughout the network.

(f) Bottleneck Layers To enhance computational efficiency, bottleneck layers are introduced. A 1×1 convolution is applied before the 3×3 convolution to reduce the number of input feature maps.

(g) Compression To further improve the efficiency and compactness of DenseNet, feature maps are compressed at the transition layers by reducing their number. The output of a transition layer is denoted by A_n.

Optimal tuning of DenseNet using proposed CDBO

DBO [42] algorithm is stimulated with ball-rolling, dancing, reproduction, stealing, and foraging characters of the dung beetle. It depends on local exploitation and global exploration. **CAViaR** [43] associated with DBO is effective for DR classification with great speed.

Dung beetle position encoding A search space Y is regarded as obtaining a superior solution. Here, ω reveals the parameter to study the **DenseNet**, and hence it is formulated by $Y = [1 \times \omega]$.

Fitness function The function is assessed regarding the variance among the results of and also the result of the target, which is formulated as,

$$f = \frac{1}{m} \sum_{n=1}^{m} [T_n - A_n]^2 \tag{8.15}$$

where, the targeted outcome is characterized as T_n, the result of the **DenseNet** is represented as A_n, and the entire data is signified as m. Here, CDBO achieved best solution based on the following step.

Step 1: Initialization Primarily, it initializes the location of dung beetle present in search space, which is specified beneath

$$X = \{X_1, X_2, \ldots, X_f, X_T\} \tag{8.16}$$

wherein, complete variables present are denoted as X_f, X_T implies the f^{th} solution, as well as the search space is termed as $\bar{X}$.

Step 2: Computation of objective function The goal of the operation is executed through employing Eq. (8.15) by figuring the variations among the results of the target as well as **DenseNet**.

Step 3: Rolling process of dung beetle

To process rolling, dung beetle's ball-rolling positioning is formulated subsequently,

$$x_i(t+1) = x_i(t) + \alpha \times \kappa \times x_i(t-1) + \beta \times \Delta_x \tag{8.17}$$

here, $\Delta_x = |x_i(t) - X^w|$

While a dung beetle successfully finds a new orientation, that may continue in rolling ball in backward manner. Subsequently, the ball-rolling dung beetle position is described as follows:

$$x_i(t+1) = x_i(t) + \tan(\theta)|x_i(t) - x_i(t-1)| \tag{8.18}$$

Here, θ represents the deflection angle which is associated with $[0, \Pi]$. Also, for female dung beetle laying eggs, the boundary selection approach is formulated below,

$$Lb^* = \max(X^* \times (1-R), Lb), \quad Ub^* = \min(X^* \times (1+R), Ub) \tag{8.19}$$

The dynamic and brood ball position in nature is clarified below,

$$B_i(t+1) = X^* + b_1 \times (B_i(t) - Lb^*) + b_2 \times (B_i(t) - Ub^*) \tag{8.20}$$

Ideal foraging border region is formulated by,

$$Lb^b = \max(X^b \times (1-R), Lb), \quad Ub^b = \min(X^b \times (1+R), Ub) \tag{8.21}$$

The position of minor dung beetle is elucidated below,

$$x_i(t+1) = x_i(t) + C_1 \times (x_i(t) - Lb^b) + C_2 \times (x_i(t) - Ub^b) \tag{8.22}$$

$x_i(t)$ signifies **data** position of i^{th} minor dung beetle at t^{th} iteration.

From equation (8.22),

$$\begin{aligned} x_i(t+1) &= x_i(t) + C_1 \times x_i(t) - C_1 \times Lb^b + C_2 \times x_i(t) - C_2 \times Ub^b \\ &= x_i(t)[1 + C_1 + C_2] - C_1 \times Lb^b - C_2 \times Ub^b \end{aligned}$$

Applying **CAViaR**,

$$X_i(t) = \delta_0 + \sum_{m=1}^{q} \delta_m^* \cdot x_i(t-m) + \sum_{n=1}^{r} \delta_n f(x_i(t-n)) \tag{8.23}$$

Assume $q = r = 2$

$$X_i(t) = \delta_0 + \delta_1 X_i(t-1) + \delta_2 X_i(t-2) + \delta_3 f(X_i(t-1)) + \delta_4 f(X_i(t-2)) \tag{8.24}$$

Applying equation (8.24) in equation (8.23), the updated equation of dung beetle is,

$$\begin{aligned} x_i(t+1) = [1 + C_1 + C_2][\delta_0 + \delta_1 X_i(t-1) + \delta_2 X_i(t-2) + \delta_3 f(X_i(t-1)) + \delta_4 f(X_i(t-2))] \\ - C_1 \times Lb^b - C_2 \times Ub^b \end{aligned} \tag{8.25}$$

Here, C_i denotes random number, Ub^b signifies upper bound foraging, at the vector (0,1) is implied as C_2, Lb^b represents lower bound of ideal foraging, and at $(t-2)^{\text{th}}$ iteration $X_i(t-2)$ is position of i^{th} solution, at $(t-1)^{\text{th}}$ iteration $X_i(t-1)$ is position of i^{th} solution.

Step 4: The location information

During iteration process of the thief's location, it is expressed as:

$$x_i(t+1) = X^b + S \times g \times \left(\left|x_i(t) - X^t\right| + \left|x_i(t) - X^b\right|\right) \tag{8.26}$$

Step 5: Termination

CDBO often executes a process to attain an improved solution. The algorithm presents the pseudocode of CDBO.

Algorithm 1 Pseudocode of CDBO

Input: $x_i(t)$
Output: $x_i(t+1)$
Start
Prime particle's population
Compute fitness function by employing equation (8.15)
while $t \leq T_{\max}$ **do**
    Dung beetle's ball rolling
    if $r < Q$ **then**
        By utilizing equation (8.17), update ball rolling dung beetle location
    else
        By employing equation (8.20), update ball rolling dung beetle location
    end if
    Using equation (8.25), update the global dung beetle position
    Location is updated by applying equation (8.26)
    Measure fitness value for novel location
end while
Terminate

8.3 Experimental Results

Validation of the proposed deep learning model using Training set and K-fold cross-validation:

An assessment comparison of CDBO_DenseNet along with existing methods and earlier proposed methods is performed by altering the Training set and K-values as shown in 8.1 Comparative study Table 8.1 correctly displays the results obtained by CDBO_DenseNet compared with existing methods.

Table 8.1: Comparative study of CDBO_DenseNet

Analysis based upon	**Metrics**	**CNN**	**DCNN**	**CRSLO RideNN**	**HPLBO DMN**	**Proposed CDBO DenseNet**
Training set	Accuracy	83.0	85.5	90	91.3	**94.1**
	Sensitivity	83.6	84.9	89.2	91.7	**93.9**
	Specificity	82.9	84.9	89.2	86.7	**91.7**
K-value = 9	Accuracy	78	82.5	84.4	89.7	**94**
	Sensitivity	80.3	82.3	84	97.7	**92.9**
	Specificity	78	93.7	94.8	86.9	**93.4**

Comparative Analysis

Table 8.2: Comparative discussion for segmentation accuracy

Based on	**Metrics**	Watershed Transform+ RBFNN	Lesion detection +DCNN	**SeGAN**	HPLBO_ K-Net	Proposed CDBO_DJS
Training set	Segmentation Accuracy (%)	78.8	83	85.6	88.4	92.3
	K-fold value	78	83.3	84.7	89.6	93.1

Validation results of proposed methods on low- and high-resolution images:

In this section, the efficacy of the proposed frameworks in performing DR classification is analyzed separately for various methodologies, including CNN [44], DCNN [45], CRSLO-based RideNN, HPLBO_DMN, and CDBO_DenseNet, when applied to both low-resolution and high-resolution images. Table 8.3 examines the performance of the aforementioned deep learning models in classifying diabetic retinopathy (DR) using high-resolution image with 90% training dataset. It is observed that CDBO_DenseNet achieves the highest performance across all metrics. This demonstrates its superior capability in both detecting true positives and avoiding false positives. HPLBO_DMN follows as the second-best performing model. Indicating robust but slightly less effective performance compared to CDBO_DenseNet. CNN has the lowest performance. Highlighting the limitations of basic convolutional architectures in processing high-resolution images. The results indicate that more advanced deep learning architectures like CDBO_DenseNet and HPLBO_DMN outperform simpler models (CNN, DCNN) in handling high-resolution images for classification tasks. CDBO_DenseNet, in particular, sets a benchmark for accuracy, sensitivity, and specificity, reflecting its high adaptability and precision for this use case. These findings emphasize the importance of selecting advanced models for high-resolution image classification to achieve optimal results.

Table 8.3: Comparative analysis of Deep learning techniques on high-resolution images with 90% of the training dataset.

Methodology	**CNN**	**DCNN**	CRSLO-based RideNN	**HPLBO_DMN**	**CDBO_DenseNet**
Accuracy (%)	80	82.8	86.88	90	**94**
Sensitivity (%)	79.17	82.9	86	89	**93**
Specificity (%)	79.8	82.13	86.87	89	**93**

Table 8.4: Comparative analysis of Deep learning techniques on low-resolution images with 90% of the training dataset.

Methodology	**CNN**	**DCNN**	CRSLO-based RideNN	**HPLBO_DMN**	**CDBO_DenseNet**
Accuracy (%)	78.7	81	85.7	89	**93**
Sensitivity (%)	78.9	83.27	85	89	**92**
Specificity (%)	78.67	82.17	85.16	90	**92.37**

Table 8.4 shows that **CDBO_DenseNet** demonstrates the best overall performance on low-resolution images, similar to its performance with high-resolution images. This confirms its ability to maintain high performance even with reduced image resolution. **HPLBO_DMN** is the second-best model, reflecting strong performance, though slightly lower than **CDBO_DenseNet**. **CNN** records the lowest performance, highlighting its limitations in handling low-resolution images effectively. **CDBO_DenseNet** remains the most robust model, showcasing high accuracy and balanced sensitivity and specificity across low-resolution images. This highlights its adaptability and effectiveness irrespective of image resolution. The overall drop in performance for all models compared to high-resolution images underlines the importance of image quality in enhancing model accuracy. However, the advanced model **CDBO_DenseNet** is less impacted by resolution changes, emphasizing its robustness and potential in resource-constrained applications.

Performance Evaluation of Low-Resolution Vs High-Resolution Image

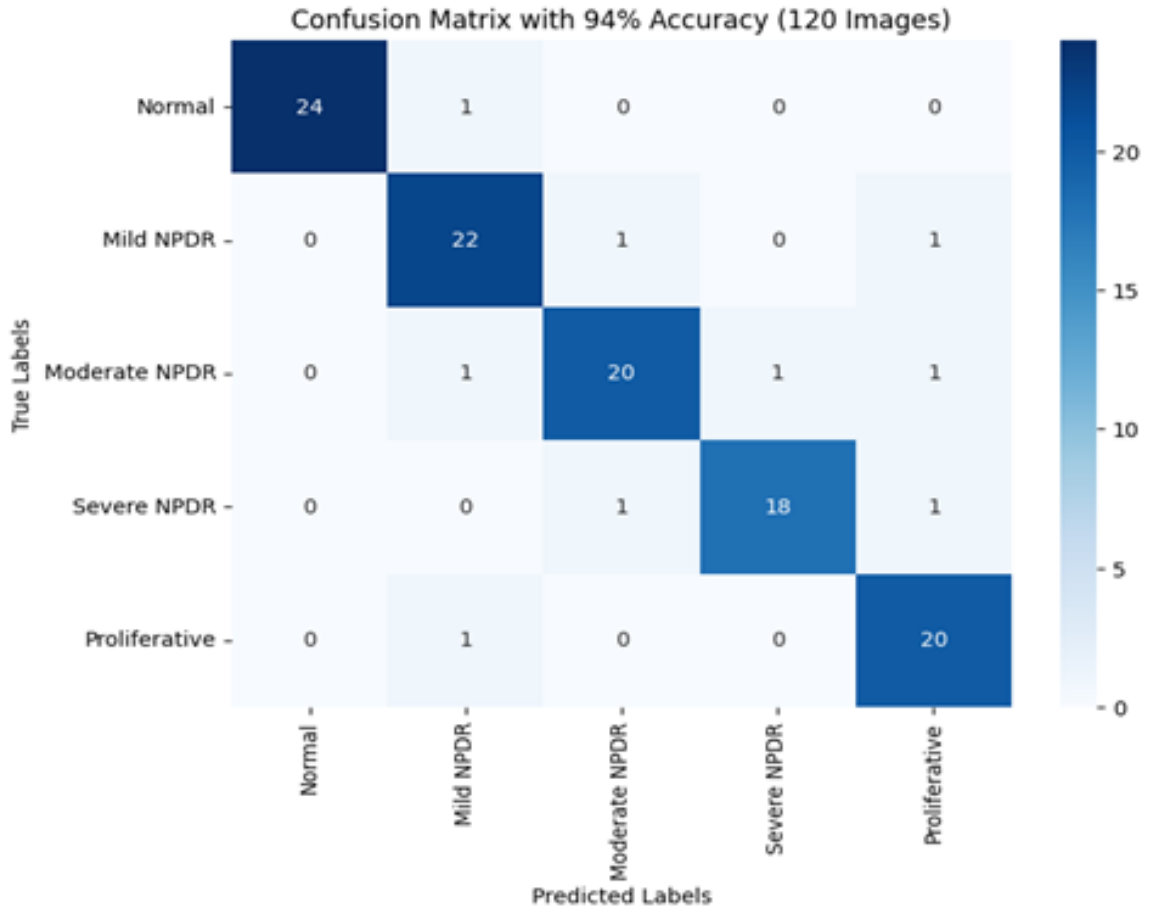

Figure 8.1: Performance analysis of CDBO_DensNet model for high-resolution images

The confusion matrix in Figure 8.1 presents the performance of the CDBO_DenseNet model on high-resolution diabetic retinopathy (DR) images, achieving an accuracy of 94%. Similarly, Figure 8.2 illustrates the model's performance on low-resolution DR images, with an accuracy of 93%. In each matrix, the diagonal values indicate correctly classified samples for each class, while the off-diagonal values represent misclassifications, where the true label differs from the predicted label. The higher accuracy for high-resolution images highlights the model's superior classification performance compared to low-resolution inputs.

Performance Evaluation of Low-Resolution Vs Low-Resolution Image

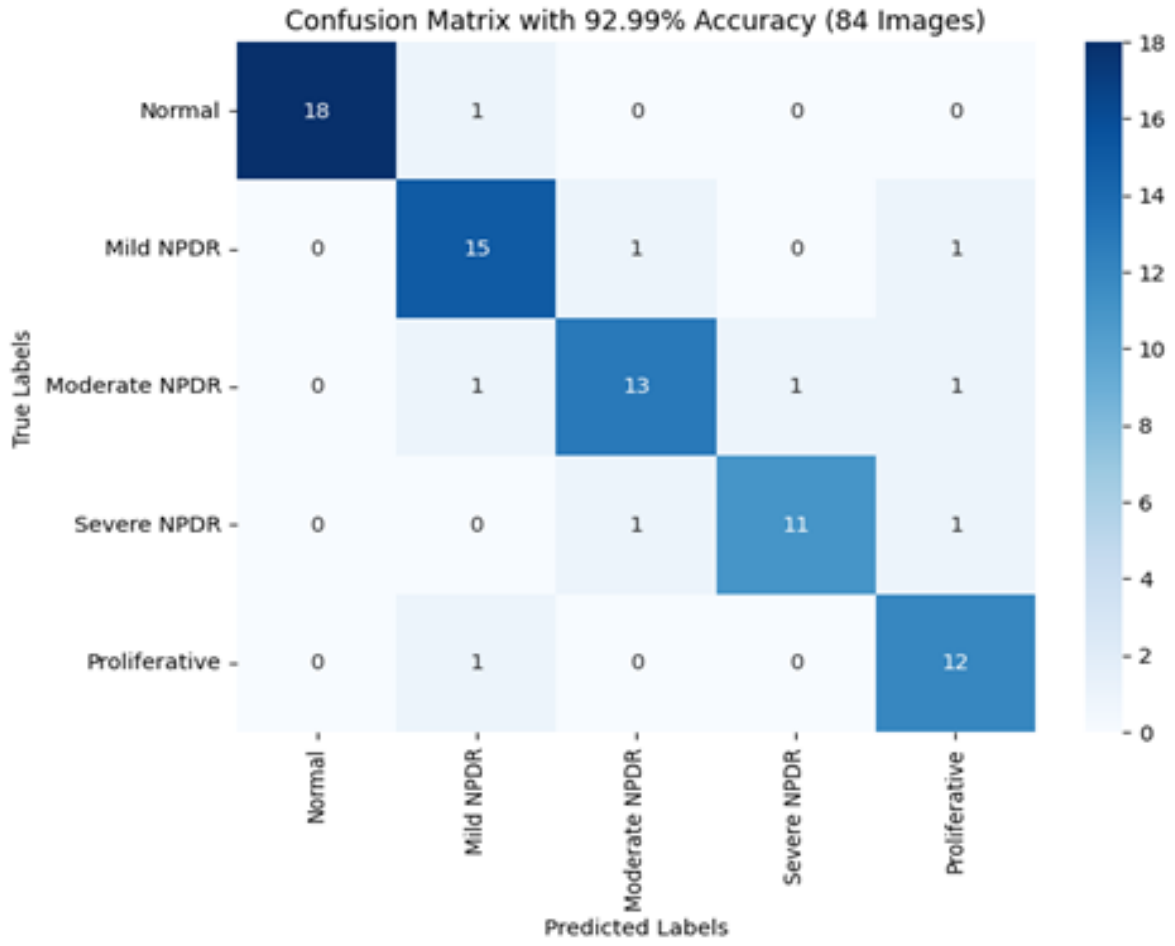

Figure 8.2: Performance analysis of CDBO_DensNet model for Low-resolution images

Loss Analysis of Individual and Hybrid Optimizers

Table 8.5 examines how the loss (a measure of error) changes as the number of K-folds in cross-validation. The data reflects the performance of three different models: Cavier_DenseNet, DBO_DenseNet, and Proposed CDBO_DenseNet.

Table 8.5: Loss vs. K-fold Cross Validation

K-fold	Cavier_DenseNet	DBO_DenseNet	Proposed CDBO_DenseNet
5	0.25	0.22	0.20
6	0.23	0.20	0.18
7	0.21	0.18	0.15
8	0.19	0.16	0.12
9	0.18	0.15	0.08

Table 8.5 highlights the performance of different DenseNet-based models across varying K-fold cross-validation settings. The proposed **CDBO_DenseNet** consistently demonstrates the lowest loss values across all K-folds, indicating superior generalization and training stability. Compared to **DBO_DenseNet** and **Cavier_DenseNet**, the loss decreases more significantly as the number of folds increases, showcasing its robustness. Notably, at 9-fold validation, the proposed model achieves a minimal loss of **0.08**, affirming its strong predictive reliability.

Table 8.6: Loss vs. Training Set Percentage

Training Set (%)	**Cavier_DenseNet**	**DBO_DenseNet**	**Proposed CDBO_DenseNet**
50	0.25	0.22	0.20
60	0.22	0.20	0.18
70	0.20	0.18	0.15
80	0.18	0.15	0.110
90	0.16	0.12	0.075

Table 8.6 shows the relationship between the size of the training set (expressed as a percentage of the total dataset) and the loss for the same three models.

The loss analysis conducted on individual and hybrid optimizers highlights the consistent and reliable performance of the hybrid optimization approach as shown in Table 8.6. Among these, the proposed CDBO_DenseNet stands out, exhibiting the lowest loss values across all training set percentages. This indicates its superior ability to minimize training error while maintaining robust generalization. The results validate the effectiveness of combining CAViaR and Dung Beetle Optimization (DBO) techniques in improving the convergence of the model, reducing overfitting, and achieving efficient optimization. These findings reinforce the model's adaptability and suitability for handling diverse and complex datasets, ensuring accurate classification in diabetic retinopathy detection tasks.

Conclusion

This research introduces the CDBO_DenseNet model for classifying Diabetic Retinopathy (DR) with several key advantages. The process begins by enhancing input images from the database using a mean filter in the pre-processing unit, which helps reduce noise and improve image clarity, making it easier to detect subtle features indicative of DR. Next, Deep Joint Segmentation is employed for lesion segmentation, offering the benefit of accurately segmenting multiple lesion types simultaneously, which enhances the reliability of the diagnosis. The AVNet model is then utilized for artery and vein classification within the enhanced images, providing the advantage of better distinguishing vascular structures that can influence DR detection. Following this, essential features are extracted, capturing critical details that contribute to a more precise DR classification.

Finally, the DenseNet model, tuned using the CDBO algorithm, is used for DR classification. The integration of CDBO allows for efficient optimization of DenseNet's parameters, improving the model's performance and robustness. As a result, the CDBO_DenseNet model achieved high accuracy, sensitivity, and specificity rates of 94%, 92.9%, and 93.4%, respectively, with 90% of the training dataset, demonstrating its effectiveness in accurately diagnosing DR even with complex datasets. Deep Joint Segmentation achieved a segmentation accuracy of 92.3% which show superior performance as compared to other methods. Also, loss analysis of individual and hybrid optimizers indicates consistent performance of the hybrid model suggesting that it is the most efficient model in terms of minimizing loss, regardless of the dataset size.

Chapter 9

DATASET

Dataset Description of the proposed Methods

The database description of the established methods is estimated depending on seven different datasets, namely DRIVE DS, STARE, Messidor, DIARETDB0, DIARETDB1, IDRID, and a Real-time Images dataset.

- **DRIVE DS Image Dataset:** Studies on retinal vascular segmentation are made possible by the DRIVE (Digital Retinal Images for Vascular Extraction) collection. Patients with diabetic retinopathy in the Netherlands provided images for the DRIVE dataset [**?**]. The dataset includes images of 400 diabetic patients aged between 25 and 90 years.

- **STARE Image Dataset:** The Structured Analysis of the Retina is also known as STARE [**?**]. The National Institutes of Health in the United States funded the project. Over its duration, thirty professionals from engineering, physics, and medicine have contributed.

- **Messidor Image Dataset:** The acronym MESSIDOR [**?**] (in French) refers to "Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology." It was created to support computer-assisted diabetic retinopathy diagnosis. The dataset is freely available for academic and scientific use.

- **DIARETDB0 Colour Fundus Dataset:** This dataset contains 130 PNG format images of color fundus photographs.

- **DIARETDB1 Colour Fundus Dataset:** This dataset includes 89 PNG images of color fundus photographs.

- **IDRID Image Dataset:** The IDRID dataset [**?**] is categorized into three parts: segmentation, disease grading, and localization.

  – Segmentation: Utilizes original color fundus and lesion ground truth images.

  - Disease Grading: Combines original images with labels for DR severity and diabetic macular edema.
  - Localization: Involves labels for optic disc and fovea center locations.

- **Real-Time Images Dataset:** This dataset comprises conventional fundus camera images and smartphone-based fundus images. Patients were recruited from the retina outpatient clinic at Somaiya Hospital, Mumbai. Ethical approval and informed consent were obtained.Patients' eyes were dilated with tropicamide (5.0 mg/mL) and phenylephrine (100 mg/mL). Imaging was done using an iPhone 6 and a 20D lens, which provides a field of view of up to 50°. The setup is lightweight (200 g) and operable with one hand. All images are PNG files and were exported for analysis.Fundus images were also captured using a conventional fundus camera (TOPCON TRC-50EX). All images underwent consistent postprocessing. A total of 125 eyes from 70 patients (mean age: 55, SD: 8.89 years) were imaged, resulting in 84 smartphone-based images and 120 conventional fundus camera images. Ground truth annotations were provided by two ophthalmologists.

Table 9.1: Summary of datasets of fundus images and their respective resolution

Dataset Name	**Image Resolution (pixels)**
DRIVE DS	768 × 584
STARE	605 × 700
Messidor	1440 × 960 to 2304 × 1536
DIARETDB0	1500 × 1152
DIARETDB1	1500 × 1152
IDRID	4288 × 2848
Smartphone-Based Fundus (SBF)	350 × 355 to 650 × 550
Fundus Camera	1936 × 1296

Chapter 10

Conclusion

Summary of Proposed Methodology and Results

Novel Methodology

- Proposed three different techniques for the classification of Diabetic Retinopathy:
  - Chronological Rider Sea Lion Optimization (CRSLO)-based Rider Optimization Algorithm based Neural Network RideNN
  - Hunter-Prey Ladybug Beetle Optimizer Deep Maxout Network (HPLBO_DMN)
  - CAViaR with Dung Beetle Optimization-enabled DenseNet (CDBO_DenseNet)
- Employed a comprehensive approach involving:
  - Preprocessing
  - Lesion segmentation
  - Artery-to-vein segmentation
  - Feature extraction

Performance

- The proposed **CDBO_DenseNet** model achieved high performance compared to the other two proposed methods, with:
  - **Accuracy:** 94%
  - **Sensitivity:** 92.9%
  - **Specificity:** 93.4%
- When evaluated on unseen low-resolution images, CDBO_DenseNet maintained robust generalization capabilities:

- **Accuracy:** 93%
- **Sensitivity:** 92%
- **Specificity:** 92.37%

Optimization Techniques

- The loss analysis of individual and hybrid optimization algorithms shows that the proposed **CDBO_DenseNet** model:
  - Consistently achieves the **lowest loss values** across all training set percentages
  - Demonstrates superior **optimization capabilities** with minimized error
  - Enhances model convergence and robustness across diverse datasets
- These results establish CDBO_DenseNet as a reliable and effective solution for diabetic retinopathy classification.

Clinical Relevance

- The proposed methods demonstrate strong potential for use in:
  - **Telemedicine** applications
  - **Remote DR diagnosis**, particularly in resource-limited settings where high-resolution imaging may not be available

Chapter 11

Future Scope

The research will be further enhanced in the future to provide support to patients with DR elaborated below: In future research, the DR will be classified using a large, infected eyes database and by utilizing Artificial Intelligence approaches. The total classes will be increased in the future to predict the classification performance of designed approaches. Moreover, new high-quality samples will be generated from the actual database in the future by utilizing Generative Adversarial Networks (GAN) in DR classification. In the future, various research investigations will be performed to enhance the lesion image contrast using an adaptive histogram equalization algorithm. Also, will develop web applications or Mobile App that integrate AI-based DR Detection can facilitate early detection and timely intervention. Using multimodal imaging techniques such as fundus images with OCT, could enhance accuracy in detecting tiny lesions.

Chapter 12

PUBLICATIONS

Publications

1. V. Sagvekar, M. Joshi, "Diabetic Retinopathy Monitoring Using Smartphone: A Review", *International Journal of Engineering Science Invention (IJESI)*, ISSN (Online): 2319-6734, ISSN (Print): 2319-6726, `www.ijesi.org`, Volume 9, Issue 3, Series I, March 2020, pp. 67–72.

2. "Early Diagnosis of Diabetic Retinopathy using Radiomic Features", presented at *IEEE Conference on Technologies for Future Cities*, 2021.

3. V. Sagvekar, M. Joshi, "Segan-Based Lesion Segmentation and Optimized RideNN for Diabetic Retinopathy Classification", *Biomedical Engineering: Applications, Basis and Communications*, Volume 35, Number 04, 2023, Article 2350008.
   DOI: 10.4015/S1016237223500084 **(Scopus Indexed Journal)**

4. Vidya Sagvekar, Manjusha Joshi, Minu Ramakrishnan, Ajay Dudani, "Hybrid Hunter-Prey Ladybug Beetle Optimization Enabled Deep Learning for Diabetic Retinopathy Classification", *Biomedical Signal Processing and Control*, Volume 95, Part B, 2024, Article 106346.
   DOI: 10.1016/j.bspc.2024.106346 **(Scopus Indexed Journal)**

5. 5) V. Sagvekar, M. Joshi, M. Ramakrishnan and A. Dudani, "CDBO_DenseNet: CAViaR Dung Beetle Optimization Enabled Deep Learning for Diabetic Retinopathy Classification," 2024 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2024, pp. 1-6, doi: 10.1109/esci59607.2024.10497412. Publisher: IEEE and Selected as Best Paper of the Session

References

[1] A. Ahmad, A. Mansoor, R. Mumtaz, M. Khan, and S. Mirza, "Image processing and classification in dr: A review," in *European Workshop on Visual Information Processing (EUVIP)*, pp. 1–6, IEEE, 2014.

[2] W. Wang and A. C. Lo, "Diabetic retinopathy: pathophysiology and treatments," *International Journal of Molecular Sciences*, vol. 19, no. 6, p. 1816, 2018.

[3] O. Faust, U. R. Acharya, E. Y. K. Ng, K. H. Ng, and J. S. Suri, "Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review," *Journal of Medical Systems*, vol. 36, pp. 145–157, 2012.

[4] V. Sagvekar and M. Joshi, "Diabetic retinopathy monitoring using smartphone: A review," *International Journal of Engineering Science Invention (IJESI)*, vol. 9, no. 3, Series I, pp. 67–72, 2020. ISSN (Print): 2319-6726, `www.ijesi.org`.

[5] "Diabetic retinopathy — our services." `https://retinams.com/our-services/diabetic-retinopathy/`. Accessed: 2025-07-18.

[6] T. Li, W. Bo, C. Hu, H. Kang, H. Liu, K. Wang, and H. Fu, "Applications of deep learning in fundus images: A review," *Medical Image Analysis*, vol. 69, p. 101971, 2021.

[7] U. Iqbal, "Smartphone fundus photography: a narrative review," *International Journal of Retina and Vitreous*, vol. 7, no. 1, p. 44, 2021.

[8] Y. Xu, Z. Zhou, X. Li, N. Zhang, M. Zhang, and P. Wei, "Ffu-net: Feature fusion u-net for lesion segmentation of diabetic retinopathy," *BioMed Research International*, 2021.

[9] Y. Guo and Y. Peng, "Carnet: Cascade attentive refine net for multi-lesion segmentation of dr images," *Complex Intelligent Systems*, vol. 8, no. 2, pp. 1681–1701, 2022.

[10] P. Saranya and S. Prabakaran, "Automatic detection of non-proliferative diabetic retinopathy in retinal fundus images using convolution neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, pp. 1–10, September 2020.

[11] T. Araújo, G. Aresta, L. Mendonça, S. Penas, C. Maia, Â. Carneiro, A. Mendonça, and A. Campilho, "Data augmentation for improving proliferative diabetic retinopathy detection in eye fundus images," *IEEE Access*, vol. 8, pp. 182462–182474, 2020.

[12] R. Rajalakshmi, R. Subashini, R. Anjana, and V. Mohan, "Automated dr detection in smartphone-based fundus photography using artificial intelligence," *Eye*, vol. 32, no. 6, pp. 1138–1144, 2018.

[13] I. Qureshi, J. Ma, and Q. Abbas, "Dr detection and stage classification in eye fundus images using active deep learning," *Multimedia Tools and Applications*, vol. 80, pp. 11691–11721, 2021.

[14] M. Karakaya and R. Hacisoftaoglu, "Comparison of smartphone-based retinal imaging systems for diabetic retinopathy detection using deep learning," *BMC Bioinformatics*, vol. 21, no. Suppl 4, p. 259, 2020.

[15] M. Wintergerst, L. Jansen, F. Holz, and R. Finger, "A novel device for smartphone-based fundus imaging and documentation in clinical practice: Comparative image analysis study," *JMIR Mhealth Uhealth*, vol. 8, no. 7, p. e17480, 2020.

[16] I. Martínez-Aguirre, Y. Dorville, J. Maroto, M. Sánchez, M. Zarzosa, and F. Benito, "Smartphones for ophthalmic fundus imaging: A review," *Revista Electrónica de Biomedicina*, vol. 3, pp. 66–73, 2017.

[17] L. Haddock, D. Kim, and S. Mukai, "Simple, inexpensive technique for high-quality smartphone fundus photography in human and animal eyes," *Journal of Ophthalmology*, 2013.

[18] L. Qiao, Y. Zhu, and H. Zhou, "Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms," *IEEE Access*, vol. 8, pp. 104292–104302, 2020.

[19] P. Scanlon, R. Malhotra, R. Greenwood, *et al.*, "Comparison of two reference standards in validating two field mydriatic digital photography as a method of screening for diabetic retinopathy," *British Journal of Ophthalmology*, vol. 87, pp. 1258–1263, 2003.

[20] W. Sun, F. Su, and L. Wang, "Improving deep neural networks with multi-layer maxout networks and a novel initialization method," *Neurocomputing*, vol. 278, pp. 34–40, 2018.

[21] D. Le, M. Alam, C. Yao, J. Lim, Y. Hsieh, R. Chan, D. Toslak, and X. Yao, "Transfer learning for automated octa detection of diabetic retinopathy," *Translational Vision Science Technology*, vol. 9, no. 2, pp. 35–35, 2020.

[22] S. Gayathri, V. P. Gopi, and P. Palanisamy, "A lightweight cnn for diabetic retinopathy classification from fundus images," *Biomedical Signal Processing and Control*, vol. 62, p. 102115, September 2020.

[23] S. Kathiresan, A. Sait, D. Gupta, S. Lakshmanaprabu, A. Khanna, and H. Pandey, "Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model," *Pattern Recognition Letters*, March 2020.

[24] F. Malerbi, R. Andrade, P. Morales, J. Stuchi, D. Lencione, J. de Paulo, M. Carvalho, F. Nunes, R. Rocha, D. Ferraz, and R. Belfort Jr, "Diabetic retinopathy screening using artificial intelligence and handheld smartphone-based retinal camera," *Journal of Diabetes Science and Technology*, vol. 16, no. 3, pp. 716–723, 2022.

[25] S. Dutta, B. Manideep, S. Basha, R. Caytiles, and N. Iyengar, "Classification of diabetic retinopathy images by using deep learning models," *International Journal of Grid and Distributed Computing*, vol. 11, no. 1, pp. 89–106, 2018.

[26] S. Honnungar, S. Mehra, and S. Joseph, "Diabetic retinopathy identification and severity classification," 2016. Unpublished or conference paper without full citation details.

[27] S. Gayathri, V. Gopi, and P. Palanisamy, "A lightweight cnn for dr classification from fundus images," *Biomedical Signal Processing and Control*, vol. 62, p. 102115, 2020.

[28] M. Mahmoud, S. Alamery, H. Fouad, A. Altinawi, and A. Youssef, "An automatic detection system of diabetic retinopathy using a hybrid inductive machine learning algorithm," *Personal and Ubiquitous Computing*, pp. 1–15, 2021.

[29] S. Renukalatha and K. Suresh, "Automatic roi extraction in noisy medical images," *ICTACT Journal on Image & Video Processing*, vol. 7, May 2017.

[30] S. Pascual, A. Bonafonte, and J. Serra, "Segan: Speech enhancement generative adversarial network," in *Proceedings of Interspeech*, March 2017.

[31] H. Kang, Y. Gao, S. Guo, X. Xu, T. Li, and K. Wang, "Avnet: A retinal artery/vein classification network with category-attention weighted fusion," *Computer Methods and Programs in Biomedicine*, vol. 195, p. 105629, 2020.

[32] D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "Bayesian filtering for location estimation," *IEEE Pervasive Computing*, vol. 2, no. 3, pp. 24–33, 2003.

[33] T. Gadekallu, N. Khare, S. Bhattacharya, S. Singh, P. Maddikunta, I. Ra, and M. Alazab, "Early detection of diabetic retinopathy using pca-firefly based deep learning model," *Electronics*, vol. 9, no. 2, p. 274, 2020.

[34] W. Zhang, J. Pang, K. Chen, and C. Loy, "K-net: Towards unified image segmentation," *Advances in Neural Information Processing Systems*, vol. 34, pp. 10326–10338, 2021.

[35] I. Naruei, F. Keynia, and A. Sabbagh Molahosseini, "Hunter–prey optimization: Algorithm and applications," *Soft Computing*, vol. 26, no. 3, pp. 1279–1314, 2022.

[36] S. Safiri and A. Nikoofard, "Ladybug beetle optimization algorithm: application for real-world problems," *The Journal of Supercomputing*, pp. 1–50, 2022.

[37] W. Sun, F. Su, and L. Wang, "Improving deep neural networks with multi-layer maxout networks and a novel initialization method," *Neurocomputing*, vol. 278, pp. 34–40, 2018.

[38] P. Tang, H. Wang, and S. Kwong, "G-ms2f: Googlenet based multi-stage feature fusion of deep cnn for scene recognition," *Neurocomputing*, vol. 225, pp. 188–197, 2017.

[39] C. Maheshan and H. Kumar, "Performance of image pre-processing filters for noise removal in transformer oil images at different temperatures," *SN Applied Sciences*, vol. 2, pp. 1–7, January 2020.

[40] K. Mahesh and J. Renjit, "Deepjoint segmentation for the classification of severity-levels of glioma tumour using multimodal mri images," *IET Image Processing*, vol. 14, no. 11, pp. 2541–2552, 2020.

[41] G. Huang, Z. Liu, L. Van Der Maaten, and K. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

[42] J. Xue and B. Shen, "Dung beetle optimizer: A new meta-heuristic algorithm for global optimization," *The Journal of Supercomputing*, vol. 79, no. 7, pp. 7305–7336, 2023.

[43] R. Engle and S. Manganelli, "Caviar: Conditional autoregressive value at risk by regression quantiles," *Journal of Business & Economic Statistics*, vol. 22, no. 4, pp. 367–381, 2004.

[44] P. Saranya and S. Prabakaran, "Automatic detection of non-proliferative diabetic retinopathy in retinal fundus images using convolution neural network," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–10, 2020.

[45] X. Luo, W. Wang, Y. Xu, Z. Lai, X. Jin, B. Zhang, and D. Zhang, "A deep convolutional neural network for diabetic retinopathy detection via mining local and long-range dependence," *CAAI Transactions on Intelligence Technology*, 2023.

[46] R. Liu, X. Wang, Q. Wu, L. Dai, X. Fang, T. Yan, J. Son, S. Tang, J. Li, Z. Gao, A. Galdran, J. M. Poorneshwaran, H. Liu, J. Wang, Y. Chen, P. Porwal, G. S. Wei Tan, X. Yang, C. Dai, H. Song, M. Chen, H. Li, W. Jia, D. Shen, B. Sheng, and P. Zhang, "Deepdrid: Diabetic retinopathy-grading and image quality estimation challenge," *Patterns*, vol. 3, p. 100512, May 2022.

[47] R. Bernardes, P. Serranho, and C. Lobo, "Digital ocular fundus imaging: a review," *Ophthalmologica*, vol. 226, no. 4, pp. 161–181, 2011.

[48] K. Alabdulwahhab, W. Sami, T. Mehmood, S. Meo, T. Alasbali, and F. Alwadani, "Automated detection of diabetic retinopathy using machine learning classifiers," *European Review for Medical and Pharmacological Sciences*, vol. 25, no. 2, 2021.

[49] L. Math and R. Fatima, "Adaptive machine learning classification for diabetic retinopathy," *Multimedia Tools and Applications*, vol. 80, no. 4, pp. 5173–5186, 2021.

[50] J. Amin, M. Sharif, M. Yasmin, H. Ali, and S. Fernandes, "A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions," *Journal of Computational Science*, vol. 19, pp. 153–164, 2017.

[51] S. Goel, S. Gupta, A. Panwar, S. Kumar, M. Verma, S. Bourouis, and M. Ullah, "Deep learning approach for stages of severity classification in dr using color fundus retinal images," *Mathematical Problems in Engineering*, pp. 1–8, 2021.

[52] N. Prabhu, D. Bhoir, and N. Shanbhag, "Diabetic retinopathy screening using machine learning for hierarchical classification," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 10, pp. 1943–1948, 2019.

[53] D. Gunasekeran, D. Ting, G. Tan, and T. Wong, "Artificial intelligence for diabetic retinopathy screening, prediction and management," *Current Opinion in Ophthalmology*, vol. 31, no. 5, pp. 357–365, 2020.

[54] S. Natarajan, A. Jain, R. Krishnan, A. Rogye, and S. Sivaprasad, "Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial intelligence system on a smartphone," *JAMA Ophthalmology*, vol. 137, no. 10, pp. 1182–1188, 2019.

[55] M. Hagos and S. Kant, "Transfer learning-based detection of diabetic retinopathy from small dataset," *arXiv preprint arXiv:1905.07203*, 2019.

[56] H. Kälviäinen, H. Uusitalo, J. Järvinen, J. Pärkkä, and J. Kämäräinen, "Diaretdb1 diabetic retinopathy database and evaluation protocol," in *Medical Image Understanding and Analysis*, p. 61, Citeseer, 2007.

[57] K. Koo and E. Cha, "Image recognition performance enhancements using image normalization," *Human-centric Computing and Information Sciences*, vol. 7, no. 1, pp. 1–11, 2017.

[58] Z. Ying, G. Li, Y. Ren, R. Wang, and W. Wang, "A new low-light image enhancement algorithm using camera response model," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 3015–3022, 2017.

[59] A. Longo, S. Morscher, J. Najafabadi, D. Jüstel, C. Zakian, and V. Ntziachristos, "Assessment of hessian-based frangi vesselness filter in optoacoustic imaging," *Photoacoustics*, vol. 20, p. 100200, 2020.

[60] X. Yang, X. Shen, J. Long, and H. Chen, "An improved median-based otsu image thresholding algorithm," *AASRI Procedia*, vol. 3, pp. 468–473, 2012.

[61] N. Zulpe and V. Pawar, "Glcm textural features for brain tumor classification," *International Journal of Computer Science Issues (IJCSI)*, vol. 9, no. 3, p. 354, 2012.

[62] A. Tartar, N. Kilic, and A. Akan, "Classification of pulmonary nodules by using hybrid features," *Computational and Mathematical Methods in Medicine*, 2013.

[63] S. Mirjalili, S. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.

[64] A. Azar, H. Elshazly, A. Hassanien, and A. Elkorany, "A random forest classifier for lymph diseases," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 2, pp. 465–473, 2014.

[65] S. Kumar, A. Adarsh, B. Kumar, and A. Singh, "An automated early diabetic retinopathy detection through improved blood vessel and optic disc segmentation," *Optics & Laser Technology*, vol. 121, p. 105815, 2020.

[66] P. Khojasteh, B. Aliahmad, and D. Kumar, "Fundus images analysis using deep features for detection of exudates, hemorrhages and microaneurysms," *BMC Ophthalmology*, vol. 18, no. 1, pp. 1–1, 2018.

[67] M. Parul and D. Rohil, "Optimized image steganography using discrete wavelet transform (dwt)," *International Journal of Recent Development in Engineering and Technology*, vol. 2, pp. 75–81, February 2014.

[68] V. M. Mane and D. Jadhav, "Holoentropy enabled-decision tree for automatic classification of diabetic retinopathy using retinal fundus images," *Biomedical Engineering Biomedizinische Technik*, vol. 62, no. 3, pp. 321–332, 2017.

[69] J. Wu and C. Liu, "Finger-vein pattern identification using svm and neural network technique," *Expert Systems with Applications*, vol. 38, pp. 14284–14289, October 2011.

[70] L. D. Hubbard, R. J. Brothers, W. N. King, L. X. Clegg, R. Klein, L. S. Cooper, A. R. Sharrett, M. D. Davis, J. Cai, and A. R. in Communities Study Group, "Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study," *Ophthalmology*, vol. 106, no. 12, pp. 2269–2280, 1999.

[71] S. Sergyan, "Color histogram features based image classification in content-based image retrieval systems," in *Proceedings of the 6th International Symposium on Applied Machine Intelligence and Informatics*, pp. 221–224, 2008.

[72] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition," in *Tenth IEEE International Conference on Computer Vision (ICCV)*, vol. 1, pp. 786–791, 2005.

[73] D. Binu and B. S. Kariyappa, "Ridenn: A new rider optimization algorithm-based neural network for fault diagnosis in analog circuits," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 1, pp. 2–26, 2018.

[74] R. Masadeh, B. Mahafzah, and A. Sharieh, "Sea lion optimization algorithm," *Sea*, vol. 10, no. 5, 2019.

[75] IEEE Dataport, "Indian diabetic retinopathy image dataset (idrid)." `https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid`. Accessed: January 2022.

[76] A. Silik, M. Noori, W. Altabey, R. Ghiasi, and Z. Wu, "Comparative analysis of wavelet transform for time-frequency analysis and transient localization in structural health monitoring," *Structural Durability & Health Monitoring*, vol. 15, no. 1, p. 1, 2021.

[77] H. Kang, Y. Gao, S. Guo, X. Xu, T. Li, and K. Wang, "Avnet: A retinal artery/vein classification network with category-attention weighted fusion," *Computer Methods and Programs in Biomedicine*, vol. 195, p. 105629, 2020.

[78] K. Fan and T. Hung, "A novel local pattern descriptor—local vector pattern in high-order derivative space for face recognition," *IEEE Transactions on Image Processing*, vol. 23, no. 7, pp. 2877–2891, 2014.

[79] "Drive ds-20 images dataset." https://datasets.activeloop.ai/docs/ml/datasets/drive-dataset/. Accessed: May 2023.

[80] "Stare dataset." https://www.kaggle.com/datasets/vidheeshnacode/stare-dataset. Accessed: May 2023.

[81] "Messidor dataset." https://www.adcis.net/en/third-party/messidor/. Accessed: May 2023.

[82] L. Xiong, L. Zhang, X. Huang, X. Yang, W. Huang, H. Zeng, and H. Tang, "Dcast: a spatiotemporal model with densenet and gru based on attention mechanism," *Mathematical Problems in Engineering*, pp. 1–12, 2021.

[83] M. Ceylan and H. Yaşar, "A novel approach for automatic blood vessel extraction in retinal images: Complex ripplet-i transform and complex valued artificial neural network," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 24, no. 4, pp. 3212–3227, 2016.

[84] C. Köse, U. Şevik, C. İkibaş, and H. Erdöl, "Simple methods for segmentation and measurement of diabetic retinopathy lesions in retinal fundus images," *Computer Methods and Programs in Biomedicine*, vol. 107, no. 2, pp. 274–293, 2012.